AUGUST
1962
VOLUME 26
NUMBER 4

JOURNAL OF CONSULTING PSYCHOLOGY

AMERICAN PSYCHOLOGICAL ASSOCIATION

JOURNAL OF CONSULTING PSYCHOLOGY

EDWARD S. BORDIN, Editor
University of Michigan

BOYD MCCANDLESS, Associate Editor
University of Iowa

Advisory Editors: John W. Atkinson, *University of Michigan*. Frank X. Barron, *University of California, Berkeley*. John M. Butler, *University of Chicago*. Allen T. Dittmann, *National Institute of Mental Health*. Seymour Fisher, *Baylor University*. Donald W. Fiske, *University of Chicago*. Harrison G. Gough, *University of California, Berkeley*. William L. Hays, *University of Michigan*. Jerome Kagan, *Fels Research Institute*. Paul E. Meehl, *University of Minnesota*. Daniel R. Miller, *University of Michigan*. Alan K. Rosenwald, *University of Illinois*. Saul Rosenzweig, *Washington University*. Julian B. Rotter, *Ohio State University*. Nevitt Sanford, *University of California, Berkeley*. Seymour B. Sarason, *Yale University*. Robert R. Sears, *Stanford University*. Edwin S. Shneidman, *Veterans Administration, Los Angeles, California*. Read D. Tuddenham, *University of California, Berkeley*. George S. Welsh, *University of North Carolina*.

The *Journal of Consulting Psychology* is the clinical journal of the American Psychological Association. It is devoted primarily to the publication of original research relevant to psychological diagnosis, psychotherapy and counseling, personality, and the dynamics of behavior. Case studies, theoretical contributions, descriptions of clinical techniques, and discussions of the professional problems of clinical psychology also appear in its pages.

Manuscripts submitted for *regular* or *early* publication and manuscripts for Brief Reports should be sent to the Editor: Edward S. Bordin, Department of Psychology, University of Michigan; Ann Arbor, Michigan. All manuscripts should be submitted *in duplicate* to facilitate referral to the Advisory Editors.

Style of Manuscripts. Manuscripts must be prepared in the style described in the *Publication Manual of the American Psychological Association (1957 Revision)*, obtainable for $1.00 from the Washington office of the APA.

Early Publication. Authors may secure the early publication of an acceptable article in the first issue to go to press after its acceptance by arranging for the article to be printed as extra pages, increasing the number of pages that subscribers receive. Information about charges for early publication may be obtained from the Managing Editor.

American Documentation Institute. To save printing costs, materials supplementary to articles may be deposited with the ADI. To obtain copies, order by Document Number as cited in the reference footnote, from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. The indicated fee must be remitted in advance. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Reprints. Authors receive fifty reprints gratis. Additional quantities may be ordered when returning proofs.

Helen Orr
Managing Editor

Sharon Stone
Editorial Assistant

Communications concerning subscriptions, change of address, claims for the nonreceipt of a number, advertising, and other business matters, should be sent to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D.C. Address changes must reach the subscription office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

The *Journal of Consulting Psychology* is published bimonthly in February, April, June, August, October, and December, the yearly volume comprising approximately 576 pages. The subscription per year is $10.00, foreign $10.50, single number $2.00.

Published by the American Psychological Association, Inc.
at Prince and Lemon Sts., Lancaster, Pa.
and 1333 Sixteenth St. N.W., Washington, D.C.

Second-class postage paid at Lancaster, Pa.
© 1962 by the American Psychological Association, Inc.

JOURNAL OF CONSULTING PSYCHOLOGY

August 1962 Vol. 26, No. 4

Contents

ARTICLES

Rusty: A Brief Behavior Disorder: *William K. Boardman* - - - - - - - - - - - 293
Punishment Revisited: *Albert Bandura* - - - - - - - - - - - - - - - - - 298
On the Definition of Problems and the Interpretation of Symptoms: *Daniel R. Miller* 302
Wechsler-Bellevue Comparisons in Groups with "Chronic" and "Current" Lateralized and Diffuse Brain Lesions: *Kathleen B. Fitzhugh, Loren C. Fitzhugh, and Ralph M. Reitan* - 306
Ego Strength and Anxiety in Discrimination Conflict Performance: *Lenard Worell and Larry K. Hill* - 311
Personality and Hypnotic Susceptibility: *Peter J. Lang and A. David Lazovik* - - - 317
The Influence of Symbolic Stimulus Value on Perceived Size in Chronic Schizophrenia: *Valadimir Pishkin, Thomas E. Smith, and H. W. Leibowitz* - - - - - - - - - - 323
Lack of Correlation between the California Psychological Inventory and Hypnotic Susceptibility: *Ernest R. Hilgard and Lillian W. Lauer* - - - - - - - - - - - 331
Intellectual Impairment in Schizophrenia: *Richard M. Griffith, Betsy Worth Estes, and Selwyn A. Zerof* - 336
The Effect of Order of Presentation on Perception of TAT Pictures: *Adelaide Dollin and James M. Sakoda* - 340
Relationships among Three Measures of Conflict over Hostility: *J. Thor Nelson and Seymour Epstein* - 345
Anxiety and the Spiral Aftereffect Test: *James H. Bryan and Edward Loder* - - - - 351
Perceptual and Cognitive Functioning of Retarded Readers: *Richard H. Walters and Helen Doan* - 355
Zen Buddhism: A Psychological Review: *Edward W. Maupin* - - - - - - - - - 362

NOTES AND COMMENT

Transparency in the Edwards Personal Preference Schedule: *Maurice Korman and Frances Coltharp* - 379
Naysaying and the California Psychological Inventory: *David Rosenhan* - - - - - 382
Social Desirability and the Relative Validities of Achievement Scales: *Alfred B. Heilbrun, Jr.* - 383
Bender-Gestalt Figure Rotations in the Mentally Retarded: *A. B. Silverstein and Philip J. Mohan* - 386

BRIEF REPORTS

Observable Signs of Anxiety or Distress during Psychotherapy: *Allen Raskin* - - - 389
A Further Investigation of the Relation between Anxiety and Digit Symbol Performance: *Robert A Johnston and Herbert J. Cross* - - - - - - - - - - - - - 390
A Note on the Validity of the Wechsler Adult Intelligence Scale for Mental Retardates: *Gary M. Fisher* - 391
The Effect of Repeated Testing on Anxiety Scale Scores: *A. W. Bendig and Gail Bruder* 392
Race of Examiner and Patient in Psychiatric Diagnosis and Recommendations: *Elaine Dorfman and Robert J. Kleiner* - 393
Personality Characteristics (EPPS), Level of Expectation, and Performance: *Carroll E. Izard* - 394

Brief Reports

The *Journal of Consulting Psychology* will accept Brief Reports of research studies in clinical psychology for early publication without expense to the author. The procedure is intended to permit the publication of soundly designed studies of specialized interest or limited importance which cannot now be accepted because of lack of space. Several pages in each issue will be devoted to Brief Reports, published in the order of their receipt without respect to the dates of receipt of the regular articles. Most Brief Reports appear in the first or second issue to go to press following their final acceptance.

An author who wishes to submit a Brief Report:

1. Sends the Brief Report, limited to one printed page and prepared according to the specification given below.

2. Also sends to the Editor a full report of the research study, in sufficient detail to give a clear account of its background, procedure, results, and conclusions, which will be filed with the American Documentation Institute to insure indefinite availability.

3. Prepares at least 100 mimeographed copies of the full report, which the author will send without charge to all who request it as long as the supply lasts.

4. Agrees not to submit the full report to another journal of general circulation.

Specifications

Brief Report. The Brief Report should give a clear, condensed summary of the procedure of the study and as full an account of the results as space permits.

To insure that the Brief Report will be no longer than one printed page, its typescript, including all matter except the title and the author's lines, must not exceed 85 lines averaging 42 characters and spaces in length. Set the typewriter margins for short lines of 42 characters, which are 3.5 inches long in elite typing, and 4.2 inches long in pica.

The manuscript of the Brief Report must be *double spaced* throughout. Except for its short lines, it follows the standard style of the 1957 revision of the *APA Publication Manual.* Headings, tables, and references are avoided or, if essential, must be counted in the 85 lines. Each Brief Report must be accompanied by a footnote in the style below, which is typed on a separate sheet and *not* counted in the 85-line quota:

[1] An extended report of this study may be obtained without charge from John Doe (giving the author's full name and address) or for a fee from the American Documentation Institute. Order Document No. —— from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress; Washington 25, D. C., remitting in advance $— for microfilm or $— for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

Extended report. Because the extended report is intended for photoduplication, and is not copy to be sent to a printer, its style should differ in several ways from that of other manuscripts: (*a*) The extended report should be typed with single spacing for economy in duplication. (*b*) Tables and figures should be placed adjacent to the text which refers to them. A caption should be typed below each figure. (*c*) Footnotes should be typed at the bottom of the page on which reference is made to them. In other respects, the full report is prepared in the style specified by the *Publication Manual.*

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 293–297

RUSTY:

A BRIEF BEHAVIOR DISORDER [1]

WILLIAM K. BOARDMAN

College of Medicine, Baylor University

When a child who is a behavior problem is seen for possible treatment, it frequently can be concluded that the relationship between the child and a parent or parent-surrogate is an unhealthy one for the child and has resulted in a behavior disorder. In the majority of such cases there is a good chance that psychotherapy can be used to eliminate the disorder. If the child and at least the most psychodynamically significant adult can be worked with, over time the desired changes in interpersonal relationships may be achieved.

The "standard" psychotherapeutic approach has proven to be the most effective in the great majority of such cases. Maximum control over the behavioral and emotional expressions of parents and child and their interrelationships is possible. Progress—even without insight—receives regular support and thus has a better chance of being maintained. New information and psychodynamic changes may be integrated into the psychotherapy. Finally, the therapist is in the best position to judge accurately the point at which sufficient progress has been made to terminate treatment.

There is an alternative approach which is effective with certain behavior disorders. It is essentially a "short cut" involving the application of simple learning theory principles in circumstances where complex psychodynamic factors can be considered to be of minimal concern. Its effectiveness depends upon several conditions: (*a*) The misbehavior in question should be of recent origin or well circumscribed and must be voluntary (this rules out enuresis, tics, stuttering, and various other "neurotic" behaviors). (*b*) It can be assumed that the misbehavior produces a great deal of direct emotional gratification and does not represent to any important degree the disguised expression of repressed affect which requires catharsis through further symptomatic expression. (*c*) It is possible to arrange for desired alternative behavior to lead to immediate rewards, while important gratifications can be withheld during periods of misbehavior. (*d*) It is possible to apply the procedure on a relatively continuous uninterrupted basis. (*e*) The child and the adults involved all appear to have a good potential for rapid behavior change.

The following case illustrates this approach. The history of Rusty's behavior disorder is taken from an account furnished by his mother.

Rusty's behavior problems started when he was 5 years 7 months old and I went into the hospital for surgery. There were complications, I had four operations, and was confined in the hospital for 9 weeks. In the daytime, Rusty and his three sisters, ages 10, 8, and 3½, were left in the care of a maid, who has been with me for 8 years, and at night with their grandmother until their father came home from work, usually around 9:00 P.M.

Rusty's first symptoms of being upset were wetting the bed at night, taking money out of his father's wallet, and riding his bike back and forth across a very busy thoroughfare.

(The maid and grandmother could not control Rusty. His grandmother would simply report his misbehavior to his father after Rusty had gone to bed. Climbing on the roof of the house and hammering all the knots out of a neighbor's decorative wood fence were typical of his misdeeds.)

I came home from the hospital and assumed that things would return to normal. Rusty's behavior seemed to get worse, rather than better. He defied me at every opportunity, continued to cross the thoroughfare on his bicycle, and even took matches and scorched furniture in his bedroom one day.

My husband and I decided that the best thing would be to enroll Rusty in a private school in the

[1] Editor's Note. *JCP* is committed to a policy of fostering the processes of discovery as well as those of verification. This is the reason for publishing W. K. Boardman's case report which seems to challenge a variety of existing conceptions of psychotherapy. Challenges demand replies, and A. Bandura and D. R. Miller, whose comments follow, were chosen to reply with the expectation that their remarks would come from theoretical commitments different from each other and from W. K. Boardman's.

fall where he could be in the first grade. By public school requirements he was not eligible for first grade because of his birthday. We entered him in a private school and thought that this would help solve our problem because his time would be occupied and he wouldn't have an opportunity to do all of the things he had been doing. The first 3 days of school went beautifully.

The fourth day, Rusty walked out of school and walked over 3 miles home. He told the teacher that he had to go to the bathroom and instead just left school. He was stopped by two men from the school who were out looking for him and asked his name. He gave a false one, so the men let him go. He later told me that he knew the men were from school and that if he told them his right name that they would take him back. He said that he didn't like school and didn't want to get taken back. I did take him back when he got home, though. When we got back to school the teacher began to tell him (in front of me) that a mad dog could have gotten him and torn him to pieces, or that a snake could have jumped out of the weeds and bitten him, and if that didn't happen then the police would get him, put him in jail, keep him on bread and water, make him go to school, and would whip him every time he gave a wrong answer. In spite of this frightening talk, Rusty stayed in school the rest of the day, and also returned to school on the following day where he remained all day. I did visit his room for about 30 minutes on this day.

The next day, Rusty reluctantly went to school. At 9:30 A.M. I received a phone call from my son at school. He said he had a stomach ache. I told him he would feel better later and that I could not come to school because I had nobody to leave the baby with. Around noon I decided to call the school to see if Rusty were still in class. I was informed that he had spent most of the morning in the office trying to call me. They wouldn't let Rusty call, but pretended to call the number for him and kept telling him that the line was busy. Well, about 1:30 P.M. Rusty left school, then went up to a house and asked if he might use the phone. He then called home himself, and finding that I wasn't at home he hung up. He then told the lady who lived there that he had been riding on a bicycle with another boy and fell off, and that the little boy went off and left him and that now he was going to have to walk home. The lady told him it was too far to walk and offered to drive him home. Rusty pointed out three different houses where he told her he lived. She would let him out at the house and wait for him to go in. Instead, he would walk up to the door of each house, look to see if she was watching, but would not go in. So, the lady would take him on to the next house where he said he lived. This went on until finally Rusty showed her the right house. I arrived home just as Rusty did and learned the story from the housewife.

That night the teacher telephoned me and asked how Rusty was. Then she proceeded to tell me that she had told Rusty that day that if he didn't start acting better at school, that his mother might go back to the hospital. That same night we reached a decision—this was not the right school for our son, and this was not the right teacher. So, the next day we registered him for kindergarten in public school (the afternoon session).

He attended kindergarten on Friday. On Monday, he left school after being there only 5 minutes. He had told the teacher he had to go to the bathroom, and when she went to get an older child to stay with her class while she took him, he ran out. This time, though, he did not return home until after 3 o'clock when he saw the other children coming home from school. He knew that I would take him right back to school if he came home.

Tuesday, Rusty was told to stay in the house all morning because he had been leaving school. He was told that when he went to school, then he would be allowed to play with other children. He ran out the front door. Two hours later, he was brought home by a policeman. Rusty had been on his way to his grandmother's house, and was only about two blocks away. He had walked about 2 miles. At this point, he got tired so he went up to a house, rang the bell, and asked to use the telephone. He then called his grandmother to come and pick him up and take him to her house, but she was not at home. The police were then summoned. First he directed the policeman to his grandmother's house and told him that this was his house and that his mother had sent him to the store and told him that if she wasn't at home when he got back he was to go to his friends' house on another street. The policeman told him that he couldn't do that, he had to take him home. So, then Rusty said he would go to his grandmother's house. He brought the officer to his own home, telling him that this was grandmother's house. (Rusty tried to keep the police officer from finding out where he lived, just as he had done previously with the housewife. The officer let him out at a corner and surreptitiously followed him home, telling his mother the circumstances.)

Later that day we had our only direct contact with Rusty and his parents. His mother related the above events, and both parents were concerned with the dangerous nature of some of his antics. From this history, the case appeared to fulfill the necessary conditions described above. Rusty's hostility was open and generalized, with little indication of repression. In fact, he appeared quite skillful in manipulating others in ways which directly satisfied his needs. His parents had not developed rigid patterns of coping with Rusty, and were willing to follow any suggestions which would alleviate the situation. They could thus be used as the agents for accomplishing behavioral change by instructing

them in detail as to how to carry out procedures in which Rusty would always be given a choice between unacceptable behavior and clearly defined acceptable behavior and would be punished severely and continuously while he persisted in undesirable acts and rewarded promptly and adequately when he shifted to appropriate behavior. While in such cases it may be assumed generally that a permissive therapeutic relationship unnecessarily prolongs acting-out, an additional factor was also involved here. Some of Rusty's behavior constituted a realistic danger to his physical safety. Permissiveness would allow this danger to continue, and he would even have opportunity to devise *new* forms of dangerous behavior.

While a secretary watched Rusty to prevent his slipping away—he made one half-hearted attempt to walk off down the hall—his parents were told that what they needed most in order to prevent such misbehavior from continuing was courage. They were instructed to buy a screen door latch for Rusty's bedroom door to use, when necessary, to keep him in his room when he had been sent there as a disciplinary measure. Spankings were suggested as appropriate only if they produced the desired results—they had been ineffective up to that point. We pointed out that, unless they literally imprisoned Rusty in his room, he had the power to run away at any time he chose. That since it was impractical from every point of view to make Rusty a permanent prisoner, they must admit to his ability to leave the home or the school, but that they should make the consequences of such behavior so drastic that he would decide for himself to be cooperative. We instructed them to tell Rusty that, if he ran away, he would be locked out of the house and would not be let in until he apologized and promised to behave. We further instructed them to refuse to give Rusty his meals while he misbehaved if it became necessary to do so in order to impress upon him that he could not flaunt authority and enjoy family privileges. They were urged to report regularly by telephone on the progress of their efforts.

Such a regimen appears harsh, but it has a distinct advantage. It is highly unlikely that a parent would not have ambivalent feelings toward a child manifesting a severe behavior disorder. With verbal or other symbolic forms of discipline, ambivalence may be expressed to a degree which seriously reduces or even reverses the intended effect of parental action. When reward and punishment are kept on more basic physical level, there is a much higher degree of efficiency in the learning situation. Both the child *and* the parents understand the rewards and punishments, and are able to recognize and profit from positive changes when they occur.

We spoke to Rusty for only a few minutes. He at first acted the part of a frightened, mistreated, misunderstood child, backing into a corner, hiding his face in his arms, and whimpering. He would not talk. (This was less than 3 hours after he had successfully manipulated a strange housewife and a police officer to suit his purposes.) We pulled him onto our knee and began by commenting that he must be very angry to act the way he had, and that we would guess that he was mad at his mother. We then suggested that his misbehavior was a lot of trouble for him and that it probably wasn't worth the effort. At this point he slid off our lap and began to talk in a very unconcerned manner about something trivial which had happened the day before. We terminated the interview and the family left. His mother reported that as soon as they were in their car, Rusty commented to his parents that they must be pretty mixed up to need to go to a doctor to find out how to make him behave. No attempt had been made to elicit important content during the few minutes spent with Rusty, but it was concluded that this child's ability to shift his behavior and expressions of affect so rapidly and dramatically indicated a minimal loss of emotional control and a good potential for learning from direct interaction with his parents. His mother's account continues.

Wednesday morning Rusty attempted to run away from home, but his father caught him and locked him in his room. He proceeded to try to rip out the window screen and escape. All implements which he could use for this purpose were removed from his room, but he continued to work at the screen with

his belt buckle. In late morning, he was let out so that he could be fed and dressed for school. He immediately darted into the bathroom and locked the door, where he stayed for about an hour before the door could be opened. He refused lunch and was dressed by force and carried by his father to school, kicking and screaming, and deposited in the classroom.

The teacher closed the door and Rusty jumped out of the window and left school. He remained away from home just as before until after 3 o'clock. He was locked out of the house because of this until after 5:30 that evening.

At 5:30 he came to the door and apologized. He was let in, fed, and allowed to follow a normal routine through the evening.

Thursday and Friday he stayed in school. Then Friday, after supper (he ate a very large meal) Rusty asked if he could go out and play. When told he couldn't, he left anyway. When he came back and found the door locked (which he had been told before he left would happen) he became pretty upset. He beat on the door with a broom, rang the doorbell, and when he found out this did him no good, he went to the fuse box and proceeded to push all the trippers, thereby turning off all the lights except the one in the garage, where he began playing. He also pushed down the lever which turned off the air conditioner. (His mother telephoned during this period and was quite upset. She was advised to take the other children and leave home for an hour, but she could not bring herself to do so.) He finally went to a neighbor's house and asked to use the phone. He called his daddy at work and told him that he was locked out. His daddy advised him to come apologize, and when he did he was let back into the house.

Saturday (This was Rusty's sixth birthday. His party and presents had been cancelled because of his bad behavior. His mother first greeted him in the morning by singing "Happy Birthday to You." This obviously ambivalent act threw Rusty into a violent fit of anger and precipitated a behavioral crisis which was the turning point in his relationship with his mother.) I told him he had to stay in his room. I locked him in, so he then cut his screen and was going to get out a window. I told him if he wanted out I would let him out the front door, but that he could not get back in. He started to walk away from the house (without socks and shoes). In about 5 minutes he was back to get his bicycle, but I locked the garage so he couldn't get it. So, he went and turned off the air conditioner at the fuse box. (The box had been padlocked, but he broke the lock.) When I told him I didn't care, he then said that he was going to climb the pole and cut the wires at the pole. He then climbed on the roof and remained there while I got in the car with the other children and drove off. When we returned about an hour later, he was playing in the backyard. (At noon his mother called and reported that she could not go on. Her husband was contacted and told that he was to go home at once and back up his wife. He was instructed to take his wife's side in the battle and to tell Rusty that any more destruction of property would be met by him with extremely severe punishment. Rusty did not come in during the 3 hours his father was at home, but neither did he make a pest of himself. At this point he had missed breakfast and lunch.) He had been given two deadlines—to come in by 12 o'clock in order to receive his lunch and 4 o'clock in order to receive his supper. He was also told by his father that when he came in he would have to spend the rest of the day in his room. He chose to stay outside until 6:15 P.M. at which time it began to rain. When he came in, he was spanked, bathed, and put in bed without supper.

This was putting it mildly. After her husband returned to work, Rusty's mother began to feel really angry for the first time. Whether she was most angry at Rusty, her husband, or the author is a moot question, but when Rusty finally came in she exploded, and he got the whipping of his life. This episode, 4 days after the family's consultation, was the turning point in overcoming Rusty's misbehavior. His mother was learning that she was capable of continuing disciplinary pressure long enough to force Rusty to choose to cooperate, but that she must not initiate hostile, rejecting behavior or it would drive him to further rebellion. Rusty was learning that rebellion, while exciting, was too costly in terms of creature comforts when he went beyond certain limits. The effectiveness of the change in relationship between Rusty and his mother is seen in her final account.

Rusty was an angel all day Sunday and most of Monday. Monday evening while at the dinner table, Rusty and his sister began to get rather wild—giggling and talking—everything except eating. So, they were warned that if they weren't finished by a certain time, their dinner would be thrown out and they would get nothing until breakfast. They continued to misbehave, so the food was thrown out. Well, Rusty was furious. He pushed the table, and in doing so turned over some coffee. He was told to go to his room. Instead he ran and locked himself in the bathroom. He was ignored. In about 15 minutes he came out and went to his room. A few minutes later when I checked on him, I found him sound asleep on the floor.

This was Rusty's final gesture of rebellion. It occurred 6 days after the family's consultation. During the following 11 months, Rusty

has remained a spirited but not a rebellious boy. He was tested with the Wechsler Intelligence Scale for Children and found to be of bright-normal intelligence.

A procedure such as this produces rapid results—or none at all. In properly selected cases the probability of success is high, and the major risk involved lies in the assumption that the parent or parents are capable of initiative, rapid learning, and insight when faced with the responsibility for all of the actions required for repairing the relationship with their child. Many parents deserve such a vote of confidence.

(Received July 20, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 298-301

PUNISHMENT REVISITED

ALBERT BANDURA

Stanford University

The development and selection of psychotherapeutic procedures are to a large extent determined by one's model of psychopathology. Until recently, the prevailing theories of behavior disorders have employed the symptom-underlying disease model in which behavior deviations are considered to be derivatives or symptoms of a host of underlying pathological psychodynamic forces. Consequently, therapeutic attention is focused not on the deviant behavior itself, but on internal processes in the form of unconscious agents, complexes, and intrapsychic conflicts.

In the last several years, a number of writers have proposed a social learning model of psychopathology that not only differs fundamentally from the psychodynamic-disease model in the conceptualization of the acquisition and maintenance of maladaptive behavior, but also generates radically different procedures for the modification of deviant behavior (Bandura, 1961; Bandura & Walters, in press; Eysenck, 1960; Ferster, 1961; Skinner, 1956; Wolpe, 1958). According to the social learning theory, so-called symptomatic behavior is viewed not as emotional disease manifestations but as learned reactions which can be modified directly by the provision of appropriate social models, and by the manipulation of response-reinforcement contingencies. Once the maladaptive behavior is altered, it is unnecessary to modify or remove an underlying pathology.

In his paper, "Rusty: A Brief Behavior Disorder," Boardman proposes essentially a two-factor theory of psychopathology, a position that is shared by many psychotherapists. On the one hand, according to Boardman, there exists a large class of psychodynamic disorders which can be treated successfully only by "standard" psychotherapeutic procedures; on the other hand, there is a restricted class of behavior disorders in which psychodynamic factors are absent and are therefore directly modifiable by procedures derived from learning theory. This dichotomous thesis implies that deviant behavior is governed by two different sets of principles (i.e., learning and psychodynamic principles) requiring fundamentally different psychotherapeutic strategies.

We have speculated elsewhere (Bandura & Walters, in press) that the necessity for invoking psychodynamic explanatory constructs may be due largely to a lack of knowledge of the genesis of maladaptive behavior. When the actual social learning history is known, psychodynamic explanations become superfluous. This point can be well illustrated by Watson and Rayner's case of Albert (1920), in whom they induced a severe phobia by several pairings of a rat and a loud sound. On the basis of stimulus generalization, the phobic reaction generalized to a wide range of furry objects including other animals, fur coats, wool, cotton, and human air. If Albert were studied in a modern psychiatric clinic, it is highly improbable that the diagnostician would write a single-sentence diagnostic report to the effect that Albert's phobia is an outgrowth of a pairing of a rat and a loud noise, and that the fear elicited by this pairing had generalized to other furry objects. It is much more likely that the diagnostician would assume that the phobia represented a symptomatic manifestation of some form of psychodynamic agents, the description of which would require at least a full page.

That the one and the same case of deviant behavior can be viewed as a complex psychodynamic disorder and as a straightforward learned reaction is exemplified in recent reinterpretations of Freud's famous case of Little Hans (Bandura & Walters, in press; Wolpe & Rachman, 1960). Clearly, the validity and utility of differentiations between psychodynamic and nonpsychodynamic behavior disorders remains to be demonstrated empirically.

Before commenting on Boardman's program of therapy for Rusty, it may be of value to speculate briefly on the factors re-

sponsible for the development and maintenance of Rusty's deviant behavior.

Children are likely to exhibit wide individual differences in response to stress produced by the abrupt and relatively prolonged separation from the mother. Aggression, dependency, withdrawal, regression, and somatization are only some of the more common modes of stress reactions that have been observed. The separation induced stress undoubtedly accounts for Rusty's heightened reactivity during the mother's hospitalization. It fails to explain, however, why aggressive disrupting reactions, rather than some other form of deviant behavior, were elicited. Evidently, Rusty's maladaptive behavior had been shaped by conditions that existed prior to the separation experience.

Although little is known about Rusty's earlier social learning history, the available anecdotal evidence suggests that his deviant behavior was largely a product of inadequate positive reinforcement. Caretaking was delegated to a maid during the daytime who, in turn, delivered Rusty to his grandmother for the evening; the father appeared to be relatively unavailable even during the period of the mother's hospitalization. In such an atmosphere of social deprivation, prosocial reactions and even mild forms of antisocial behavior are likely to go unnoticed and nonreinforced, while more intense and aversive forms of behavior force attentive and caretaking behavior from parental surrogates. If the more forceful successive responses are differentially reinforced, aggressive disrupting behavior eventually becomes strongly established. Thus, Rusty began, among other things, to climb on rooftops, dismantle neighborhood property, and ride his bicycle on a busy thoroughfare.

While Boardman interpreted Rusty's behavior as manifestations of unrepressed hostility, an alternative explanation, and the one favored by the present writer, is that Rusty's deviant responses were maintained by intermittent positive reinforcement, i.e., his reckless behavior produced aversive stimulation for his caretakers who were forced to terminate the fear provoking behavior by attending to him. If, in fact, hostility was the maintaining mechanism, the use of punishment as a treatment procedure is clearly contraindicated since punitiveness would be expected to generate and augment already existing hostile feelings and thus enhance the very behavior the treatment was designed to control.

The decision to enroll Rusty in a private school was clearly an attempt by the parents to escape his aversive control over their behavior. Unfortunately, the school provided even fewer positive reinforcements and substantial aversive stimulation so that Rusty was forced to engage in extreme forms of escape and avoidance behavior including somatic complaints, flight, and deception, not unusual reactions for a child caught up in a double avoidance conflict.

The treatment as outlined to Rusty's parents was based on the principle of differential reinforcement. They were to reward Rusty for socially approved behavior and to punish him promptly and severely for undesired reactions. The available case material reveals, however, that the parents resorted almost exclusively to punishment and made little or no use of positive reinforcement. Does this form of treatment find support in behavior theory? The answer is clearly in the negative. The ineffectiveness of punishment in eliminating undesired responses is abundantly documented not only in research on both animal and human subjects (Azrin, 1959, 1960; Bandura & Walters, 1959; Estes, 1944; Sears, Maccoby, & Levin, 1957) but is well illustrated in the parents' anecdotal accounts of Rusty's response to the punishment regime.

In the first place, the presentation of aversive stimuli in the form of severe physical punishment gives rise to conditioned anxiety which often motivates and reinforces patterns of behavior that are even more socially undesirable than the ones that punishment was intended to suppress. Thus, at the outset Rusty exhibited active avoidance behavior designed to escape a noxious teacher. When flight proved unsuccessful, physical complaint behavior emerged as an alternative escape response but this too was nonreinforced and flight appeared once again as the dominant avoidance reaction. Although this behavior elicited delayed punishment, nevertheless it became strongly established because of its effectiveness in terminating the aversive school stimulation, in bringing the mother to the

school, and eventually in removing him permanently from the fear provoking teacher.

Under the threat of punishment, Rusty was forced to learn remarkably elaborate deceptive behavior in order to escape a host of pursuers, a possible forceful return to the school, and parental punishment for his truancy. The teacher and the school office personnel provided him excellent models of deception behavior.

What began as an approach-avoidance conflict was quickly transformed under the punishment regime into a double avoidance conflict in which Rusty began to exhibit strong avoidant behavior toward the parents who had become relatively severe punishing agents. He was now escaping from the parents to the grandmother, a consequence requiring additional physical restraint and punitiveness. This illustrates a second negative by-product of punishment—not only does it produce conditioned anxiety but also avoidance of the punishing agents which further reduces their effectiveness for social training.

Thirdly, parental use of physical punishment provides the child an aggressive model for imitation, the influence of which may counteract any temporary inhibitory effects of punishment on the performance of aggressive behavior. While the intended outcome of the parents' punishment was the suppression of Rusty's aggression, it furnished him with vivid examples of how to counteraggress toward the parents.

Finally, unless the punished responses are weakly established or punishment is extremely severe, its use tends to increase the frequency and amplitude of aggressive behavior. If Rusty exhibited responses of high magnitude prior to treatment, these were relatively mild compared to the intensities of aggression and other forms of aversive counter control generated by the punitive treatment procedures. The parents, in turn, were forced to increase the severity of punishment until eventually it reached almost physiologically intolerable intensities. It was at this point that the inhibitory effects of punishment outweighed all other influences and Rusty's deviant behavior, much of which was treatment generated, was finally suppressed.

The treatment, as administered by the parents under Boardman's direction, emphasized the suppression of deviant responses over the development and maintenance of socially desired patterns of behavior. Even a casual acquaintance with learning theory would reveal that neither of these objectives can be effectively achieved by the use of punishment. On the contrary, a learning theory based therapy would focus upon positive rather than aversive control in bringing about the desired behavioral changes. Such a program of treatment would involve the positive reinforcement of socially appropriate behavior and the removal of positive reinforces supporting the disturbing deviant responses.

Much of Rusty's disruptive behavior would undoubtedly have extinguished following the withdrawal of attention rewards. Excellent demonstrations of rapid decline in the frequency of persistent disturbing responses when attentive reinforcing reactions to the deviant behavior are withheld are provided by Williams (1959) in the treatment of severe tantrum behavior, and by Ayllon (1960; Ayllon & Michael, 1959) in the modification of various classes of psychotic behavior of long standing.

Since, however, Rusty already suffered from inadequate positive reinforcement, and some of his venturesome behavior constituted a serious physical threat to himself and to others, more active positive intervention was required. The most effective method for eliminating deviant behavior while at the same time promoting socially desired changes involves the active reinforcement of incompatible adaptive responses. Had the parents reversed the reinforcement contingencies so that Rusty secured parental interest and approval for constructive prosocial behavior, the probability that he would have continued to sit on rooftops, mar and dismantle furniture, flee the household, and play highway roulette would be substantially reduced. The control of Rusty's deviant behavior could thus have been achieved without any adverse by-products by strengthening incompatible socially appropriate patterns of behavior.

So far I have presented alternative interpretations of the functional value of Rusty's troublesome behavior, and outlined a therapeutic strategy that differs from the one described in the case material. The one point

on which I find myself in complete agreement with Boardman is in the reliance on parents as allies and agents of treatment.

A problem that invariably arises in the course of child psychotherapy centers around parental requests for advice on how to alter their children's disturbing behavior. Such requests are generally ignored, evaded, or negatively reinforced by therapists because of their antiadvice training. It is typically assumed that an emotionally disturbed interpersonal relationship lies at the root of childhood deviant behavior, or that the parents unconsciously desire and instigate the very behavior about which they complain. Thus, the children's reactions are regarded as symptoms of the parental disturbance and consequently the therapeutic attention is focused on the parents' internal dynamics and intrapsychic conflicts. While it is true that parents frequently reward their children's troublesome behavior, the parental schedules of reinforcement may be regulated more by pain avoidance, as I have outlined in earlier sections of this paper, than by vicarious positive reinforcement derived from the children's reactions.

If a child's deviant patterns are to be altered to any significant degree, the parents must be informed, in one way or another, about (*a*) the cues eliciting the child's undesired behavior, (*b*) the response-reinforcement contingencies maintaining that behavior, and (*c*) the manner in which the reinforcement contingencies must be modified if clearly specified therapeutic objectives are to be achieved.

Since a psychotherapist's contact with a family is extremely limited—usually to 50 minutes per week—he can best spend the brief time available instructing the parents on how to implement the principles of behavioral change. The efficacy of this type of therapeutic approach has been demonstrated in the work of Williams and Ayllon to which reference has been made, and by Phillips (1960) in a comparative study of depth and behavioristically oriented psychotherapeutic procedures.

At times parents may be unable to carry out the necessary treatment program because of conditions influencing their own behavior. In Rusty's case, for example, the mother's medical condition and the father's relative absence from the home substantially reduce the frequency with which they can provide positive reinforcement for Rusty's prosocial behavioral repertoire. It is likely, however, that such limiting factors can be more effectively overcome by modifying manipulable variables of which the parents' behavior is a function than by delving into the parents' social learning history. In any event, it is the parents who must create and sustain the necessary learning conditions, and serve as the therapeutic agents.

REFERENCES

AYLLON, T. Some behavioral problems associated with eating in chronic schizophrenic patients. Paper read at American Psychological Association, Chicago, September 1960.

AYLLON, T., & MICHAEL, J. The psychiatric nurse as a behavioral engineer. *J. exp. Anal. Behav.*, 1959, **2**, 323–334.

AZRIN, N. H. Punishment and recovery during fixed-ratio performance. *J. exp. Anal. Behav.*, 1959, **2**, 301–305.

AZRIN, N. H. Effects of punishment intensity during variable-interval reinforcement. *J. exp. Anal. Behav.*, 1960, **3**, 123–142.

BANDURA, A. Psychotherapy as a learning process. *Psychol. Bull.*, 1961, **58**, 143–159.

BANDURA, A., & WALTERS, R. H. *Adolescent aggression*. New York: Ronald, 1959.

BANDURA, A., & WALTERS, R. H. *Deviant response patterns: The learning of childhood aggression and withdrawal*. New York: Holt, Rinehart & Winston, in press.

ESTES, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, **57**(3, Whole No. 263).

EYSENCK, H. J. (Ed.) *Behaviour therapy and the neuroses*. Oxford: Pergamon, 1960.

FERSTER, C. B. Positive reinforcement and behavioral deficits of autistic children. *Child Develpm.*, 1961, **32**, 437–456.

PHILLIPS, E. L. Parent-child psychotherapy: A follow-up study comparing two techniques. *J. Psychol.*, 1960, **49**, 195–202.

SEARS, R. R., MACCOBY, ELEANOR E., & LEVIN, H. *Patterns of child rearing*. Evanston: Row, Peterson, 1957.

SKINNER, B. F. What is psychotic behavior? In, *Theory and treatment of psychosis: Some newer aspects*. St. Louis: Washington Univer. Press, 1956.

WATSON, J. B., & RAYNER, ROSALIE. Conditioned emotional reactions. *J. exp. Psychol.*, 1920, **3**, 1–14.

WILLIAMS, C. D. The elimination of tantrum behaviors by extinction procedures. *J. abnorm. soc. Psychol.*, 1959, **59**, 269.

WOLPE, J. *Psychotherapy by reciprocal inhibition*. Stanford: Stanford Univer. Press, 1958.

WOLPE, J., & RACHMAN, S. Psychoanalytic "evidence": A critique based on Freud's case of Little Hans. *J. nerv. ment. Dis.*, 1960, **131**, 135–148.

(Received April 2, 1962)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 302–305

ON THE DEFINITION OF PROBLEMS AND THE INTERPRETATION OF SYMPTOMS

DANIEL R. MILLER

University of Michigan

In explaining any phenomenon the first step a man must take is to define the questions he intends to answer in his explanation. By choosing certain questions from among the many possible alternatives he delimits considerably the range of possible answers. Of course he tries to select those topics that he anticipates will be most significant. But how is he to know what is important before he has collected information about the phenomenon? He can only employ the leads available to him at the time. The defining of problems, then, whether it is being done by a clinician, experimenter, or poet, entails, of necessity, considerable intuition and the use of premises which can only be justified by the investigator's faith in their ultimate fruitfulness.

Disagreements among clinicians probably refer more to this first step than to any other. "What he says makes sense," is a common complaint, "but he does not seem to realize that . . ." and then the critic raises issues which, in view of his particular articles of faith and his intuitions, he deems to be more significant than those emphasized by the speaker. Such were the kinds of disagreements I found myself having with Boardman when I read his interesting and eminently reasonable report on Rusty, disagreements based on differences in fundamental positions which must be taken by anyone who attempts to understand the origins of a patient's symptoms.

Most of my differences with Boardman could be derived from disagreements about five basic positions, which will provide the outline for the remainder of this paper. We differ in our attitudes toward parsimonious explanations, the importance we attach to the goal of explaining all the presented facts as integral parts of a total organization, our concern with understanding the patient's experience, our predilection for theories of learning and ways of applying them, and our criteria for evaluating the patient's progress. Various clinicians' positions or to all five issues probably form two clusters which are most commonly labeled as "tough-minded" and "tenderminded."

According to Boardman, it is often unnecessary to invoke "complex psychodynamic factors" to explain the origin of a child's symptoms, in which case the psychotherapist can cure them quite readily by applying the "simple principles" of learning theory. But how does he rule out psychodynamics? Boardman's answer seems to be determined in great part by a bias in favor of parsimonious explanations. Apparently assuming that he does not have to worry about psychodynamics unless evidence of their existence leaps to the eye, he makes the most obvious interpretations of Rusty's plight: that he is being deliberately hostile, that he is gratified by his antisocial behavior, that his symptoms are not the indirect expression of repressed impulses, and that the parents seem sufficiently cooperative and flexible to assist the therapist in overcoming Rusty's problems.

Such conclusions reflect Boardman's willingness to accept the facts reported by the mother at their face value. He does try to satisfy one of his criteria for treatment by physical rewards and punishments by stating there is little indication of repression. In apparent justification of this claim, he adds that the boy was skillful in manipulating people in ways that satisfied his needs.

Some readers will feel that obvious interpretations are incomplete and may even be misleading if information is not collected about some additional problems. Wetting the bed at night, for example, a symptom mentioned in passing, requires further investigation because it violates the criterion that the response be voluntary. Another observation that needs to be clarified is the connection

between the onset of symptoms and the mother's absence from the home for 9 weeks, during which time she had four operations. Why should these events have stirred up hostility rather than other manifestations of disturbance? And does the author not have to account for Rusty's practice of riding his bicycle back and forth across a busy street, an apparently suicidal act which might have expressed a profound guilt? And is it not important to understand why a little boy seemed impelled to destroy himself when his mother became seriously ill and was taken to the hospital?

Questions of the type just raised are important because they initiate lines of investigation which are quite different from those discussed by Boardman. But to what extent would a group of critics who read the account of Rusty's problem agree in their listing of important unanswered questions? They might agree about some core questions, and would probably disagree about others. There are many "schools" of therapy and many questions can be raised. There being relatively little consensus about criteria for choosing among the formulations of questions the clinician must develop a high "tolerance of ambiguity."

The second major basis of disagreement concerns the goal of explaining all the information available about a patient as part of a meaningful configuration. Congruent with an inclination to favor parsimonious explanations is the tendency to select a few salient facts and to ignore others which do not seem relevant to one's impressions. The last paragraph contained some examples of facts which Boardman did not analyze in relation to Rusty's symptoms. There is much more information which Boardman reports but does not consider in his decisions about treatment. Also the facts that Rusty was spending almost all of his time with females, that some members of his family were not consistent in their treatment of his destructive acts, that Rusty had temper tantrums in which he lost all control of himself, that he used a stomach ache as an excuse for leaving school, that he expressed contempt for his parents because they need professional assistance to control him.

At the opposite extreme from Boardman are the readers who would require that all the observations and interpretations be related to each other so as to create an internally consistent total picture. Such a requirement would be based on the assumption that all of a person's behavior is comprehensible and that one way of testing the validity of interpretations is to see how they fit the molar configuration.

Probably the most debatable decision underlying the phrasing of questions concerns the importance the observer attaches to the patient's conscious and unconscious reactions. Again Boardman's position is quite extreme. He derives his conclusions almost exclusively from reports of Rusty's behavior and offers very little information about Rusty's conscious or unconcious experiences. The account concentrates on observable action and the treatment consists of observable action with clear consequences; few details of the report would have had to be altered had the protagonist been a preverbal child or even a primate.

How did the problem look to Rusty? Of the cited facts that might bear on this question, among the most intriguing are the observations that he acted like "a frightened, mistreated, misunderstood child, backing into a corner, hiding his face in his arms, and whimpering. He would not talk." The apparent meaning of these reactions is rejected, however, with the comments that Rusty was acting a part and that this behavior occurred less than 3 hours after he had sucessfully manipulated a strange housewife and a police officer to suit his purposes.

The relative stress one places on the patient's version of events is inseparable from one's theoretical biases. Boardman is obviously partial to "simple learning theory," which leads him to recommend clear physical reward of desired action and harsh physical punishment of forbidden action. He mentions the topic of psychodynamics only to dismiss it as irrelevant to the case.[1]

Readers with more faith in or concern

[1] It should be mentioned, parenthetically, that experimental evidence does not support Boardman's assumptions that harsh punishment extinguishes undesirable responses.

with psychodynamic theory would be tempted to postulate many relationships not mentioned in the paper. Why, one might ask, did Rusty become so upset when his mother had to go to the hospital? An obvious answer is that he suddenly got much less attention than when she was home. He was one of four children who had to be cared for by a busy grandmother and a father who did not quit work until late in the evening. Why did the boy have such temper tantrums? No one could restrict him after the mother went to the hospital, but it also seems probable that he had been treated inconsistently even before she became sick; otherwise he would have had more control. Questions about discipline raise new topics for investigation: why the family could not set adequate limits for Rusty's behavior; the nature of the parents' marriage; the reasons why the father was away so much.

Given a concern with psychodynamics, it would be difficult to overlook the possible effects on Rusty of spending almost all of his time with females. How is a boy almost 6 years of age to develop a masculine identity under such conditions? How can he resolve his oedipal conflict if he lacks access to adult men? Possibly Rusty is like a number of children described in the clinical literature whose hostile destructiveness represents one of the few ways in which they can display masculine initiative. Possibly the persistence of destructiveness and rebelliousness after the mother came home was "overdetermined": the symptoms insured her attention and helped Rusty to express his budding masculinity. There may even have been a third reason why the symptoms were rewarding: they dragged Rusty's father, a potential source of affection and model of masculine identification, back into the family.

Boardman might claim the clinching argument with respect to a final position about which clinicians often disagree: the basis for judging changes in the patient's behavior. The interpretation of Rusty's problems must have been correct, his therapist might say, because the treatment was successful; Rusty's rebellion terminated and he remains a "spirited" boy. But an opposing faction who reject a clinician's diagnosis would also reject his explanations why harsh punishment and strong rewards have "cured" his patient.

How might the change in Rusty's behavior be explained in terms of another theoretical position? In trying to understand Rusty's retirement from the field of battle the clinician partial to psychodynamics would first review the initial problems. His list might include such obvious symptoms as enuresis, stealing, destruction of property, running away from school, and leaving home. Concern with dynamics would also prompt him to postulate a number of internal problems. I shall assume here that evidence could be accumulated to test the validity of these impressions. The defense mechanism of turning against the self might be inferred from Rusty's suicidal manner of riding the bicycle. From the extreme fear that Rusty showed in his interview one might infer a predilection for the mechanism of projection. From his reactions to school one might also conclude that he is inclined to withdraw from places where he suffers considerably anxiety, and one might postulate tendencies toward denial and omnipotence from such incidents as his taunting his parents because they had to seek professional help in order to control him. Finally he must have had considerable difficulty in integrating the infantile, feminine, destructive, and masculine parts of his identity.

In explaining the change in Rusty's behavior most clinicians would be inclined to invoke positive reinforcement, but observers would differ in their accounts of the sources of reinforcement and the tendencies that were reinforced. Boardman stresses the parents' deliberate intentional reinforcement of appropriate behavior, but omits many unintentional sources of gratification they may have provided the boy. One was harsh punishment which, while it is not recommended for purposes of extinguishing responses, may have contributed to Rusty's altered behavior by helping him to atone for his guilt. Another was the increased attention the parents paid to their son. Previously they had sent him to a private school where his "time would be occupied"; now the father was home quite often and the mother was paying much more attention to her son. Finally, the availability of his father provided the son with a mas-

culine model for identification and may have made it possible for him to take initiative in a socially acceptable way.

Which problems were not changed by the treatment? The author does not tell whether the bed-wetting and stealing diminished and it is possible they did not. The clinician who is concerned with the structure of his patient's personality as well as his behavior must also ask whether some of the underlying sources of pathology have been changed by the treatment. In the account of Rusty's case there is nothing to indicate change in his defenses of denying problems, running away from the sources of difficulty, or attacking himself. With reference to his methods of resolving conflict, then the treatment would be considered superficial even though it eliminated the major antisocial symptom, and the prediction would be made that Rusty will fare no better than he did before when he faces his next serious emergency.

In summary, faith and personal predilections have considerable influence on the phrasing of clinical problems and the judgment of treatment. Because of the lack of objective criteria it is possible for two discussants of the same paper, such as the report on Rusty, to take different points of view, each person feeling confident that his interpretations are substantially correct. This indeterminacy is one of the most perplexing problems facing the clinician today. Of course, its seriousness has been overdrawn here in order to highlight its characteristics. A child's improvement is not always as difficult to understand as it is in Rusty's case, and it is often possible to compare the success of individual clinicians with different theoretical persuasions. In the future, psychotherapists will continue to employ intuition as a major methodological tool, but it is only a matter of time until systematic research on the learning and changing of behavior provides a sufficient number of tested principles applicable to the human adult to enable the study of therapy to become a branch of the science of behavior.

(Received April 19, 1962)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 306-310

WECHSLER–BELLEVUE COMPARISONS IN GROUPS WITH "CHRONIC" AND "CURRENT" LATERALIZED AND DIFFUSE BRAIN LESIONS[1]

KATHLEEN B. FITZHUGH, LOREN C. FITZHUGH

New Castle State Hospital, Indiana

AND RALPH M. REITAN

Indiana University Medical Center

Several studies have reported differential impairment on the Verbal and Performance parts of the Wechsler-Bellevue Scale in groups of subjects with lesions of the left as compared to the right cerebral hemisphere (Andersen, 1950; Balthazar & Morrison, 1961; Kløve, 1959a; and Reitan, 1955). Additional investigations have shown similar results when using independent variables having only inferential significance with respect to lateralization of cerebral lesions, such as the presence or absence of dysphasia (Reitan, 1960), dysphasia versus inability in copying simple geometric figures (Kløve & Reitan, 1958), right versus left homonymous visual field defects (Doehring, Reitan, & Kløve, 1961), and right versus left impairment in perception of simultaneous, bilateral tactile, auditory, and visual stimulation (Kløve, 1959b). While these studies deal with Wechsler results specifically as the dependent variables, they fall in the context of a growing number of reports that lesions of the left cerebral hemisphere cause significant impairment of language and related symbolic aspects of communication whereas lesions of the right cerebral hemisphere are associated with special impairment in dealing effectively with problems of a visuospatial nature (Doehring & Reitan, 1961; Ettlinger, Warrington, & Zangwill, 1957; Hécaen, 1961; Hécaen, Penfield, Bertrand, & Malmo, 1956; Reitan, 1959; Reitan & Tarshes, 1959; and Stark, 1961).

Most of the studies cited above have used subjects with recent and rapidly developing signs and symptoms of cerebral dysfunction. A large proportion of patients with brain lesions who are studied by psychologists in clinical situations do not fall in this category, having sustained brain damage early in life or having developed chronic brain lesions that cause no marked progression of neurological signs and symptoms. The present study was performed in order to test the generality of the principle of differential impairment resulting from lateralized cerebral lesions by using a group of patients with recently developed evidence of brain damage as compared with a group having chronic and long-standing brain lesions.

[1] Supported by Research Grants B-2558 and B-1468 from the United States Public Health Service, National Institute of Neurological Diseases and Blindness.

METHOD

Subjects

Three groups of subjects with recently developed symptoms of cerebral dysfunction having current clinical significance were composed from patients tested at the Neuropsychology Laboratory at the Indiana University Medical Center. Group R_1 (damage principally of the right cerebral hemisphere) consisted of 25 subjects (21 male; 4 female), Group L_1 (left hemisphere damage) of 18 subjects (15 male; 3 female), and Group D_1 (diffuse or bilateral cerebral dysfunction) of 25 subjects (24 male; 1 female). These groups showed no significant differences in age or education. Means and standard deviations for these variables were: Group L_1, 36.89 (11.10) and 8.00 (1.78); Group R_1, 38.20 (12.00) and 7.84 (1.81); Group D_1, 37.52 (11.30) and 8.32 (2.01). The subjects with chronic brain lesions were hospitalized at the New Castle State Hospital, a state supported institution for patients with neurological problems. Group R_2 (damage predominantly of the right hemisphere) consisted of 25 subjects (12 male; 13 female), Group L_2 (damage predominantly of the left hemisphere) of 20 subjects (7 male; 13 female), and Group D_2 (diffuse or bilateral cerebral dysfunction) of 30 subjects (12 male; 18 female). These groups showed no significant differences in age or education. Means and standard deviations for these variables were: Group R_2, 33.08 (12.70) and 5.44 (3.53); Group L_2, 31.30 (10.45) and 6.00 (3.52); Group D_2,

33.63 (12.85) and 5.37 (4.01). Brain dysfunction was manifested by seizures in every patient in these groups. The median period elapsing between the time of the first noted seizure and time of psychological testing in these groups was as follows: R_2, 26 years; L_2, 23 years; and D_2, 22 years.

Comparisons of mean age and education between the groups from the two institutions yielded consistently significant differences with respect to education. The New Castle State Hospital groups were lower in education than any of the groups from the Indiana University Medical Center. None of the comparisons of age was statistically significant, although the probability of Group L_2 being younger than R_1 and D_1 was less than .10.

Classification of subjects into their proper groups was effected by using information independently obtained from a thorough neurological evaluation of each subject. These evaluations included detailed medical history, electroencephalography, neurological examination, and, in some cases, angiography, pneumoencephalography, surgeons' notes and sketches, repeated neurological examinations, cerebral blood flow studies, and autopsy findings. Diagnoses of the subjects in each group are presented in Table 1. While certain of the diagnoses imply the presence of fairly long-standing brain dysfunction in the Medical Center groups, each patient had been admitted with recently developed complaints.

Procedure

The Wechsler-Bellevue Intelligence Scale (Form I) was administered individually to each subject. All scoring was completed prior to composing the groups. For purposes of statistical comparisons, the raw scores on each variable were pooled for Groups R_1, L_1, and D_1, then ranked from poorest to best performances. These ranks were converted to normalized standard scores (T scores) for each distribution. Comparisons of means for Groups R_1, L_1, and D_1 were effected by simple analysis of variance, followed by use of Student's t test, having secured the advantage of normalized distributions and controlled variance. The number of subjects with better Verbal than Performance total weighted scores and the converse was determined in each group. Intergroup differences in this dichotomous classification were tested using chi square applied with Yates' correction for continuity. Data from Groups R_2, L_2, and D_2 were handled entirely separately but subjected to exactly the same treatment. No evaluations of differences in level between corresponding groups from the two institutions were performed because of their significant differences in education.

RESULTS

Although the analysis of data was not based on the raw scores, mean Verbal, Per-

TABLE 1

DIAGNOSTIC DISTRIBUTIONS FOR SIX BRAIN DAMAGED GROUPS

Medical Center groups	*N*	State Hospital groups	*N*
Diffuse		Diffuse	
Cerebrovascular disease	3	Convulsive Disorder due to:	
Traumatic head injury	12	Congenital brain anomaly	3
Cerebral neoplasm	5	Infectious disease	4
Cerebral atrophy	4	Trauma	4
Multiple sclerosis, acute symptoms	1	Unknown cause	19
	25		30
Left		Left	
Cerebrovascular disease	5	Convulsive Disorder due to:	
Traumatic head injury	5	Congenital brain anomaly	5
Cerebral neoplasm	7	Infectious disease	4
Cerebral atrophy	1	Trauma	8
	18	Unknown cause	3
			20
Right		Right	
Cerebrovascular Disease	12	Astrocytoma	2
Traumatic head injury	2	Convulsive Disorder due to:	
Cerebral neoplasm	8	Congenital brain anomaly	2
Cerebral abscess	3	Infectious disease	8
	25	Trauma	8
		Unknown cause	5
			25

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR WECHSLER–BELLEVUE IQ VARIABLES

Group	Test	Right		Left		Diffuse	
		M	*SD*	*M*	*SD*	*M*	*SD*
Medical Center groups	Verbal IQ	93.76	15.94	81.67	15.54	91.64	13.96
	Performance IQ	78.76	16.18	90.78	13.84	89.36	16.10
	Full Scale IQ	85.40	14.99	84.44	15.25	89.52	14.78
State Hospital groups	Verbal IQ	75.48	18.72	68.00	14.42	72.90	17.32
	Performance IQ	73.44	19.03	71.90	19.13	78.07	16.66
	Full Scale IQ	71.72	19.16	67.10	17.80	72.93	17.17

formance, and Full Scale IQ values for each group are presented in Table 2 in order to give the actual levels of intelligence quotients in familiar terms.

Figure 1 presents graphically the mean scores on each subtest of the Wechsler-Bellevue Scale for all groups used in the investigation. Group R_1 obtained higher mean scores on each Verbal subtest than did Group L_1, while this relationship of these groups was reversed on each of the Performance subtests. The general level of Group D_1 was intermediate to the other groups, contrasting with the evidence for more specific verbal or performance impairment associated with lateralized cerebral lesions. The magnitude of intergroup differences among the patients with chronic brain lesions was less pronounced. While Group L_2 had the lowest means on five of the six Verbal subtests, and Group R_2 the highest means on four of the Verbal subtests, these groups had almost identical means on the performance subtests.

Simple analysis of variance indicated that certain significant differences were present in Groups L_1, R_1, and D_1 whereas no such evi-

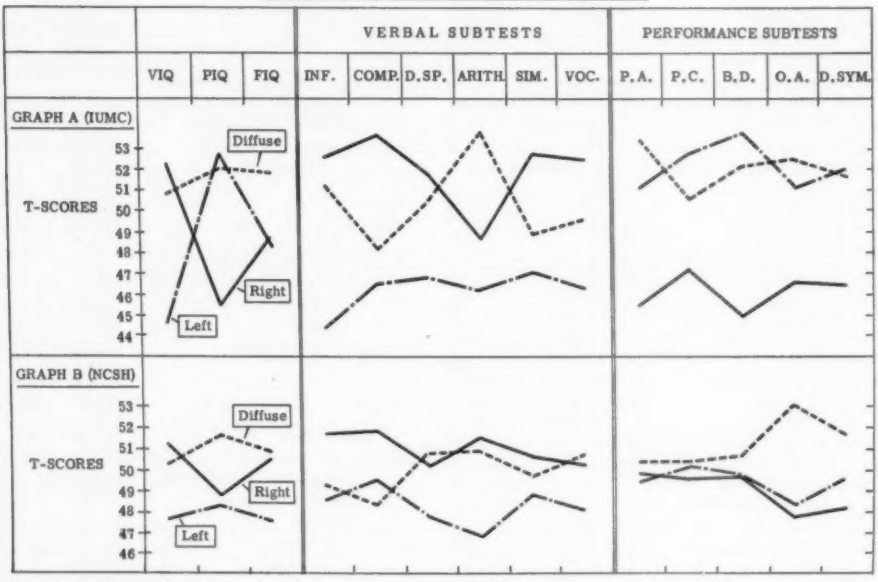

FIG. 1. Graphic presentation of mean *T* scores on Wechsler-Bellevue variables for three "current" and three "chronic" brain damaged groups.

dence occurred in the analysis of Groups L_2, R_2, and D_2. Variables showing significant intergroup variance among the patients with cerebral dysfunction of current clinical significance were Block Design ($p < .01$) and Verbal IQ, Performance IQ, Information, Comprehension, and Picture Arrangement ($p < .05$). Only these variables were considered in comparison of the three possible pairs of groups. Group L_1 obtained significantly lower scores than Group R_1 on the Verbal measures (Verbal IQ and Comprehension, $p < .05$; Information, $p < .01$) whereas Group L_1 was significantly better than Group R_1 on Block Design ($p < .01$) and the difference approached significance on Picture Arrangement ($p < .10$). The differences between Groups R_1 and D_1 occurred principally in instances in which the group with diffuse cerebral dysfunction had the higher scores: Picture Arrangement and Block Design ($p < .01$), and Performance IQ ($p < .05$). However, Group R_1 had a better mean on Comprehension than did Group D_1 ($p < .05$). The only significant differences between Groups L_1 and D_1 occurred on Verbal IQ and Information, Group D_1 having the higher values at confidence levels beyond .05.

Lower Verbal than Performance IQ values were present in 89% of Group L_1, 24% of Group R_1, and 48% of Group D_1. The difference in this dichotomous distribution between Groups L_1 and R_1 was highly significant ($p < .001$), whereas Groups L_1 and D_1 were significantly different beyond the .05 level. The difference between Groups R_1 and D_1 had a chance probability of less than .10. Lower Verbal than Performance IQ values were present in 65% of Group L_2, 44% of Group R_2, and 59% of Group D_2. While the rank-order of both sets of groups was identical with respect to these proportions, the differences among the groups with chronic brain lesions were considerably smaller and none of them approached statistical significance when tested by chi square.

DISCUSSION

Our results in groups with recently developed symptoms and signs of cerebral dysfunction are consistent with previous reports (Andersen, 1950; Kløve, 1959a; and Reitan, 1955) of differential impairment on the verbal and performance parts of the Wechsler-Bellevue Scale in groups with lateralized cerebral lesions. While a slight trend of results in the predicted direction occurred in the groups with chronic brain dysfunction, none of the comparisons showed statistically significant differences. We cannot be sure that this difference in finding is attributable solely to chronicity of brain damage, since the two sets of groups differed with respect to education and institutionalization. However, the results suggest that various factors in addition to lateralization of cerebral damage may be significant, emphasizing the complex covariance of pertinent variables within the context of "brain damage," as well as environmental conditions, in determining psychological test findings. Certainly, our results provide a basis for recommending caution with respect to application of findings obtained in one setting to clinical evaluation of patients in a different setting.

Our findings in the groups with chronic, long-standing lesions disagree to some extent with those reported by Balthazar and Morrison (1961). These investigators, using differences of specified magnitude between Verbal and Performance weighted scores as the independent variable in a group of chronic, epileptic patients, found a significant relationship with electroencephalographic classification of the patients. The extent of conflict in these studies is difficult to evaluate because of procedural differences. The present investigation inquired about the extent to which categorization of patients on the basis of extensive neurological assessment would order Wechsler-Bellevue results, whereas Balthazar and Morrison were interested in the extent to which predetermined criterion points obtained from the WISC and Forms I and II of the Wechsler scales would order electroencephalographic findings. Balthazar and Morrison proceeded to compare the Wechsler results of groups already selected for Verbal and Performance differences and concurrent agreement of electroencephalographic information. Not surprisingly, the groups selected for lower Verbal than Performance scores as well as left electroencephalographic abnormalities tended to have lower Verbal subtest scores than the

group selected for higher Verbal than Performance scores and right cerebral electroencephalographic disturbances. Balthazar and Morrison (1961) comment that their findings ". . . were similar to those reported by Kløve and Reitan and support the use of their findings with this chronically organic epileptic population" (p. 163). Their approach, however, almost certainly resulted in augmentation of intergroup mean differences and restriction of intragroup variability through use of limited ranges of verbal and performance weighted score differences as a criterion for inclusion of subjects in one group or another. Such a procedure would result in inflated *t* ratios. These considerations, together with the results of the present study, argue against uncritical transference of findings between groups which differ regarding possibly relevant characteristics of brain damage.

SUMMARY

Wechsler-Bellevue Scale results on three groups of patients with recently developed evidence of brain damage and three groups of patients with chronic, long-standing brain lesions were studied in order to test the generality of earlier reports of differential impairment on the Verbal and Performance parts of the Wechsler Scale in relation to lateralized cerebral lesions. Classification of subjects into appropriate groups was based upon detailed neurological information. Because the groups with chronic, long-standing brain damage had significantly less education than the groups with recently developed symptoms, the two sets of groups were treated independently for statistical analyses. Results for the groups with recent evidence of brain dysfunction were consistent with earlier findings of impairment on Verbal portions of the Wechsler in groups of patients with predominantly left hemisphere lesions, impairment on the Performance portion in groups with predominantly right hemisphere lesions, and an absence of differential impairment in groups with diffuse or bilateral cerebral damage. However, the groups with chronic, long-standing brain damage did not differ significantly from each other on any of the Wechsler variables, although the group with predominantly left hemisphere damage tended to be inferior to the other two chronic groups on Verbal subtests. Implications of these findings were discussed, and the authors recommended caution in applying the findings obtained in one setting to other populations which may differ in many of the characteristics relevant to the effects of brain damage.

REFERENCES

ANDERSEN, A. L. The effect of laterality localization of brain damage on Wechsler-Bellevue indices of deterioration. *J. clin. Psychol.*, 1950, **6**, 191–194.

BALTHAZAR, E. E., & MORRISON, D. H. The use of Wechsler intelligence scales as diagnostic indicators of predominant left-right and indeterminate unilateral brain damage. *J. clin. Psychol.*, 1961, **17**, 161–165.

DOEHRING, D. G., & REITAN, R. M. Language disorders in brain-damaged patients. *Arch. Neurol.*, 1961, **5**, 294–299.

DOEHRING, D. G., REITAN, R. M., & KLØVE, H. Changes in patterns of intelligence test performance associated with homonymous visual field defects. *J. nerv. ment. Dis.*, 1961, **123**, 227–233.

ETTLINGER, G., WARRINGTON, E., & ZANGWILL, O. L. A further study of visual-spatial agnosia. *Brain*, 1957, **80**, 355–361.

HÉCAEN, H. Clinical symptomatology in right and left hemispheric lesions. Paper read at Conference on Cerebral Dominance, Johns Hopkins University, Baltimore, April 1961.

HÉCAEN, H., PENFIELD, W., BERTRAND, C., & MALMO, R. The syndrome of apractognosia due to lesions of the minor cerebral hemisphere. *Arch. Neurol. Psychiat.*, 1956, **75**, 400–434.

KLØVE, H. Relationship of differential electroencephalographic patterns to distribution of Wechsler-Bellevue scores. *Neurology*, 1959, **9**, 871–876. (a)

KLØVE, H. The relationship of sensory suppression to distribution of Wechsler-Bellevue scores. Paper read at Midwestern Psychological Association, Chicago, May 1959. (b)

KLØVE, H., & REITAN, R. M. The effects of dysphasia and spatial distortion on Wechsler-Bellevue results. *Arch. Neurol. Psychiat.*, 1958, **80**, 708–713.

REITAN, R. M. Certain differential effects of left and right cerebral lesions in human adults. *J. comp. physiol. Psychol.*, 1955, **48**, 474–477.

REITAN, R. M. *The effects of brain lesions on adaptive abilities in human beings.* Indianapolis: Indiana University Medical Center, 1959.

REITAN, R. M. The significance of dysphasia for intelligence and adaptive abilities. *J. Psychol.*, 1960, **50**, 355–376.

REITAN, R. M., & TARSHES, ELAINE L. Differential effects of lateralized brain lesions on the Trail Making Test. *J. nerv. ment. Dis.*, 1959, **129**, 257–262.

STARK, ROSEMARY. An investigation of unilateral cerebral pathology with equated verbal and visual-spatial tasks. *J. abnorm. soc. Psychol.*, 1961, **62**, 282–287.

(Early publication received November 13, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 311-316

EGO STRENGTH AND ANXIETY IN DISCRIMINATION CONFLICT PERFORMANCE

LEONARD WORELL AND LARRY K. HILL[1]

Oklahoma State University

Conflict may be said to exist when the individual is confronted with the necessity of selecting between two or more simultaneously aroused incompatible response alternatives. Among the several general categories of research on conflict, the role of individual differences in response to conflict has received considerable recent attention (Block & Petersen, 1955; Broadhurst, 1957; Scodel, Ratoosh, & Minas, 1959; Worell & Castaneda, 1961a, 1961b). In line with these attempts to relate personality characteristics to conflict resolution, Korman (1960), seeking to extend the "construct validity" of the Barron Ego Strength (*Es*) scale (Barron, 1953), found support for the utility of an ego strength dimension. By exposing high and low *Es* psychiatric patients to varying degrees of brightness discrimination conflict, Korman demonstrated that high *Es* subjects resolve easy and difficult discrimination situations faster than low *Es* subjects. These findings appear to be in agreement with both clinical expectations and predictions that might be made from the relationships obtained between the *Es* scale and indices of maladjustment such as the MMPI. In this latter connection, studies by Barron (1953) and Williams and Lawrence (1954) have shown consistently high negative relations between the *Es* and most of the MMPI clinical scales, suggesting that the *Es* scale reflects a general factor of maladjustment (Barron, 1953). Thus, low *Es* subjects, being more maladjusted than high *Es* persons, might be expected to show impaired ability to resolve conflict situations.

These considerations regarding the *Es* scale are somewhat complicated, however, by the finding that the *Es* scale may also reflect motivational properties. In this connection, Ends and Page (1957) and the present authors[2] have obtained significant negative relationships between the *Es* and the Taylor Manifest Anxiety scales, suggesting that higher levels of manifest anxiety may characterize low *Es* individuals. Relevant to this characterization are the findings obtained recently by Worell and Castaneda (1961b) in an anxiety and discrimination conflict study. They found that anxious subjects, when compared to nonanxious subjects, were significantly faster in resolving low conflict situations and demonstrated a trend in the same direction under high conflict conditions. Thus, in view of the negative relationship between *Es* and manifest anxiety, it is possible that anxiety may be an important determiner of differences in conflict performance between differentiated *Es* subjects. As a result, in the following our interest is in determining first whether high and low *Es* subjects drawn from a presumably "normal" college sample perform differently under two levels of discrimination conflict, and secondly, what the effects of controlling for manifest anxiety between extreme *Es* groups are on conflict performance.

DESIGN

Two experiments were conducted to determine the relation of ego strength to the speed of conflict response. In both experiments, each subject was exposed to both high and low conflict conditions in a brightness discrimination apparatus. The simultaneous presentation of two relatively bright lights constituted the high conflict situation, while the low conflict situation consisted of the exposure of a relatively bright and dim light.

In the first experiment, high and low *Es* subjects were selected without regard to their scores on the Taylor Manifest Anxiety Scale (*MA* scale). In the second experiment, high and low *Es* groups were equated with respect to *MA* scale performance. For

[1] The authors wish to acknowledge the assistance of the Research Foundation of Oklahoma State University in conducting portions of this study.

[2] Ends and Page obtained a correlation of $-.45$ ($p < .001$) between the *Es* and the Taylor Manifest Anxiety scale; while the present authors found a correlation of $-.37$ ($p < .001$) between these scales on a sample of 354 college undergraduates.

all groups the measure of conflict performance used was the time elapsing between the presentation of the conflict condition and the initiation of a response to the conflict situation.

EXPERIMENT I

Subjects. The Barron *Es* scale, the Taylor *MA* scale, and the Lie (*L*) scale of the MMPI were administered to a total of 354 undergraduate males and females. From these, 16 subjects with scores above the 80th percentile were selected for the high *Es* group, and 16 subjects with scores below the 20th percentile were obtained for the low *Es* group. An equal number of males and females was used in each *Es* group, and no subject was included who scored above four on the *L* scale.

Apparatus. The apparatus was a modified version of that used previously (Castaneda & Worell, 1961; Worell and Castaneda, 1961b). Briefly, it consisted of two 1.5 inch in diameter, circular milk glass windows arranged 11 inches from one another on a horizontal plane and set into a black plywood panel. The illumination in each window was furnished by single 7 watt incandescent lamps capable of being varied in brightness. A 7.25 × 7.25 inch "start" platform was attached to the base of the panel equidistant from the windows, while 5 inches above each of the right and left edges of the start platform was a protruding 1-inch in diameter, circular, slightly concave push button. Lastly, a small visual ready signal was located slightly above and equidistant from the windows.

The ready signal was activated at a constant 4-second interval prior to the illumination of the windows. A Standard Electric timer was energized simultaneously with the illumination of the windows and terminated when the subject removed his hand from the start platform. Latency, consisting of the time from the onset of the illumination of the windows to the subject's raising his hand from the start platform, was recorded to the nearest 1/100 second. All reaction time scores were converted to reciprocals (1/RT × 1,000), so that all performances are presented in terms of speed of response.

TABLE 1

MEAN RESPONSE SPEEDS (1/RT × 1,000) AND STANDARD DEVIATIONS

Group	Low Conflict		High Conflict	
	M	*SD*	*M*	*SD*
Experiment I				
High *Es* Males	20.66	6.21	19.33	7.74
High *Es* Females	21.90	5.82	20.32	5.68
Low *Es* Males	17.91	3.97	15.64	3.62
Low *Es* Females	17.11	5.35	15.60	5.26
Experiment II				
High *Es* Males	17.76	5.83	16.07	7.18
High *Es* Females	22.09	6.49	20.10	6.50
Low *Es* Males	18.54	4.80	16.45	5.67
Low *Es* Females	18.29	6.59	16.96	7.23

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE FOR THE *Es* GROUPS IN EXPERIMENT I

Source	*df*	*MS*	*F*
Between	31		
Ego strength (*Es*)	1	36,635.87	3.615*
Sex (S)	1	275.27	.027
Es × S	1	1,354.70	.134
Error (b)	28	10,134.77	
Within	224		
Conflict (C)	1	6,424.62	24.746***
C × *Es*	1	110.41	.425
C × S	1	37.82	.146
C × *Es* × S	1	141.40	.545
Trials (T)	3	15,340.14	59.087***
T × *Es*	3	1,396.14	5.378**
T × S	3	230.21	.887
T × *Es* × S	3	130.89	.504
T × C	3	199.21	.767
T × C × *Es*	3	49.49	.191
T × C × S	3	203.48	.784
T × C × *Es* × S	3	214.68	.827
Error (w)	196	259.62	

* $p < .075$.
** $p < .005$.
*** $p < .001$.

Three brightness values, 2.8, 90.9, and 291.4 foot candles, were used to produce the two levels of conflict. These values were determined at each window by a MacBeth illuminometer, and the setting of variable transformers in circuit with each lamp permitted the exposure of the appropriate illuminations. The high conflict condition consisted of the two brightest stimuli (90.9 and 291.4), a difference of approximately .50 of a log unit, while the low conflict condition was comprised of the two extreme values (2.8 and 291.4), a difference of about 2.0 log units.

Procedure. Each subject, participating individually, was asked to respond to the brighter of two simultaneously exposed lights. Since previous work (Worell & Worell, in press) has suggested that the typical instructions used in RT studies may have motivational properties, the subject was informed that the experiment dealt with the accuracy of visual discrimination. At no point, did the instructions contain any indication that speed was considered important. The subject was cautioned to use his preferred hand throughout the experiment and to keep his hand upon the platform until he had decided about the location of the correct stimulus.

Each subject was presented with a total of 48

discrimination trials. Prior to the first discrimination trial, the correct "brighter" stimulus was presented alone once in each window. Both the position of the brighter light and the order of appearance of the high and low conflict trials were presented in a stratified random order. Thus, the four conditions, i.e., the two positions and the two levels of conflict, appeared an equal number of times within each block of 12 trials.

RESULTS

The mean response speeds of the high and low *Es* groups for all trials at the two levels of conflict are presented in Table 1. The results of an analysis of variance, summarized in Table 2, revealed three significant effects; the main effect of *Es* approached significance at approximately the .075 level, the main effects of conflict and trials were both significant beyond the .001 level, while the only significant interaction was that for *Es* and trial blocks ($p < .005$).

Both the trials and conflict results have been obtained many times before and require little discussion. The trials effect indicates that performance became rapid with continued practice, while the conflict finding shows that the high conflict condition produced slower speeds than the low conflict condition. Of more interest is the near significant effect of *Es*, which should, however, be evaluated in the light of the significant *Es* and trials interaction. The nature of this interaction is graphically presented in Figure 1, where one may note a progressive divergence in the response speeds of the *Es* groups over the four blocks of trials in each conflict condition. These results indicate that, although the over-all difference in performance between the high and low *Es* groups only approaches significance, high *Es* subjects increase their speed over trials at a significantly faster rate than low *Es* under both high and low conflict.

TABLE 3

MEAN RESPONSE SPEEDS PER TRIAL BLOCK

Group	1	2	3	4
		Low Conflict		
High *Es*	103.76	125.27	136.27	145.31
Low *Es*	91.56	104.79	109.40	114.48
		High Conflict		
High *Es*	90.62	113.35	129.71	142.17
Low *Es*	77.23	92.05	102.22	103.44

Note. — Above values refer to total performance over each trial block.

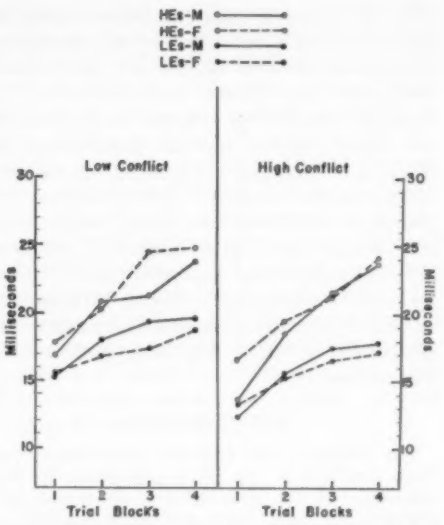

FIG. 1. Response speeds over trials for the *Es* groups in the first experiment. Each trial block represents the average performance on six trials.

In order to determine the specific meaning of the *Es* and trials interaction, a further analysis was conducted. The analysis was performed with the multiple range test (Duncan, 1955), and we required significance at or beyond the 1% level. The mean speed values per trial block for the high and low *Es* groups at the two levels of conflict are presented in Table 3. Analysis by the multiple range test ($K = 16$, $df = 100$) indicated that a difference of 18.39 between any pair of means was significant at the 1% level. An examination of the differences between high and low *Es* groups in Table 3 indicates that the only point at which the two groups do not differ significantly is during the first block of trials. Thus, under both high and low conflict conditions, extreme *Es* subjects reliably differ from one another beyond the first six trials.

DISCUSSION

Although *Es* appears to serve as an influential source of differences in response to con-

flict situations, a clear-cut interpretation of these findings is hampered by the significant association between the *Es* and *MA* scales. Reference has already been made to the negative relationship between these scales in our total sample, and an examination of Table 4 emphasizes the possible importance of this association where low *Es* subjects are shown to have had consistently higher *MA* scale scores than high *Es* subjects. In view of the highly significant nature of this difference ($p < .005$) between *Es* groups with respect to the *MA* scale, a second experiment was performed in which *MA* scale performance was controlled.

EXPERIMENT II

The procedure and apparatus were identical to those described in the initial experiment. As before, each of the high and low *Es* groups was composed of eight males and eight females. In contrast to the first experiment, however, high and low *Es* subjects, within each sex, were selected in pairs according to their *MA* scale scores. Thus, high and low *Es* males were used if their *MA* scale scores were within one point of each other; the same was the case for high and low *Es* females.

RESULTS

Before the major results with regard to conflict are presented, consideration should be given to the *MA* scale scores of the *Es* subjects. An inspection of Table 4 indicates that there are only slight differences in anxiety between the *Es* groups within each sex, but the difference between sexes is quite substantial. While statistical analysis revealed that the within sex difference was not significant, the between sex analysis showed that the females were significantly more anxious than the males ($p < .005$).

TABLE 4

MEAN *Es* AND *MA* SCALE SCORES OF SUBJECTS IN BOTH EXPERIMENTS

Group	*N*	Experiment I		Experiment II	
		Es	*MA*	*Es*	*MA*
High *Es* Males	8	54.50	10.13	55.38	13.50
Low *Es* Males	8	41.00	22.75	43.13	14.50
High *Es* Females	8	53.75	12.25	52.88	18.38
Low *Es* Females	8	36.50	24.50	38.75	18.63

TABLE 5

SUMMARY OF ANALYSIS OF VARIANCE FOR THE *Es* GROUPS IN EXPERIMENT II

Source	*df*	*MS*	*F*
Between	31		
Ego strength (*Es*)	1	4,805.53	.371
Sex (S)	1	10,724.03	.827
Es × S	1	9,456.46	.730
Error (b)	28	12,962.13	
Within	224		
Conflict (C)	1	7,247.23	27.880***
C × *Es*	1	8.78	.034
C × S	1	30.65	.118
C × *Es* × S	1	161.38	.621
Trials (T)	3	14,724.99	56.648***
T × *Es*	3	508.45	1.956
T × S	3	240.20	.924
T × *Es* × S	3	363.42	1.398
T × C	3	32.80	.126
T × C × *Es*	3	14.80	.057
T × C × S	3	173.52	.668
T × C × *Es* × S	3	108.26	.416
Error (w)	196	259.94	

*** $p < .001$.

Turning to the conflict data, the mean performances of the high and low *Es* groups for all trials at each level of conflict are contained in the lower portion of Table 1. In addition, a graphic presentation of the performance of the *Es* groups over the four trial blocks may be found in Figure 2. An analysis of variance, summarized in Table 5, indicated that the control for manifest anxiety reduced the number of significant effects to two; in particular, only the main effects of conflict and trials remained significant. The previously found near significant effect of *Es* and the highly significant trials and *Es* interaction were not obtained. Moreover, despite the divergence of the high *Es* females from the other *Es* groups evident in Figure 2, the triple interaction of trials, *Es* and sex failed to attain a significant level.

DISCUSSION

Considering the two experiments as a unit, the principal findings indicate that the effect of ego strength on conflict performance is contingent on the presence of manifest anxiety differences among *Es* individuals. High *Es*

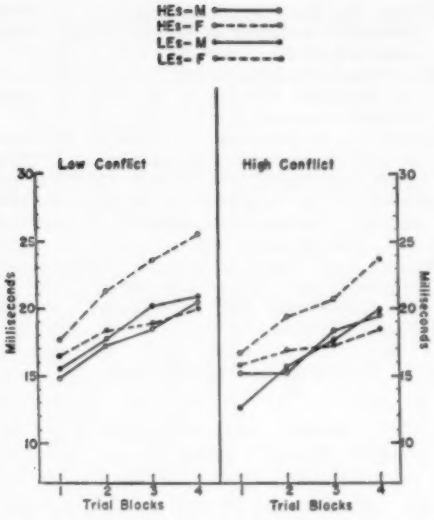

FIG. 2. Response speeds over trials for the *Es* groups in the second experiment. Each trial block represents the average performance on six trials.

groups demonstrated a consistent and, over trials, a progressively increasing superiority in performance over low *Es* subjects in both high and low conflict situations. The results with the *Es* scale agree in general with findings obtained by Korman (1960) in a similar conflict study, although the former study did not obtain a trial and conflict interaction. The performance differences here, however, were eliminated when ego strength groups were equated for manifest anxiety level.

Several aspects of the findings merit consideration. The significant interaction between ego strength and trials observed in the first experiment indicates that high *Es* subjects improve in their performance at a significantly faster rate than low *Es* subjects. This may be interpreted as signifying that ego strength contributes to differential adaptation to conflict conditions. Aside from the agreement of this finding with Barron's (1953) statement that high *Es* individuals show greater adaptability, this result is not incompatible with the more rapid improvement shown by high *Es* patients in psychotherapy (Barron, 1953). Thus, the greater progressive efficiency with which high *Es* subjects resolve experimentally aroused conflicting tendencies appears coordinate to the greater skill evidenced by such individuals in resolving more emotional conflicts. In addition, the finding of no significant performance differences between high and low *Es* groups of equivalent anxiety levels also appears to have relevance for the significant positive relationship found between ego strength and therapeutic improvement (Barron, 1953). Our result suggests that some patients with low ego strength may respond as well to psychotherapy as those with high ego strength if their anxiety levels are roughly the same. If this is the case, then the use of both the *Es* and *MA* scales might be expected to increase the predictability of therapeutic success beyond that already achieved by the *Es* scale alone.

SUMMARY

The present study was primarily oriented toward an extension of our understanding of the role of personality characteristics in the efficiency of conflict resolution. High and low ego strength subjects, as determined by the Barron Ego Strength (*Es*) scale, were exposed in two experiments to relatively high and low levels of conflict. In the first experiment, extreme *Es* groups were used without regard to scores on a scale of manifest anxiety; in the second study, extreme *Es* groups were equated with respect to manifest anxiety levels. The result indicated that when high and low *Es* subjects are selected without regard to anxiety level, high *Es* subjects resolve relatively simple and difficult discrimination conflict situations faster and with more rapid improvement over trials than low *Es* subjects. However, nonsignificant differences in conflict performance were found when the level of manifest anxiety between extreme *Es* groups was controlled. The implication of these findings for characteristics associated with the *Es* scale and psychotherapy were briefly considered.

REFERENCES

BARRON, F. An ego strength scale which predicts response to psychotherapy. *J. consult. Psychol.*, 1953, **17**, 327–333.

BLOCK, J., & PETERSEN, P. Some personality correlates of confidence, caution, and speed in a

decision situation. *J. abnorm. soc. Psychol.*, 1955, **51**, 34–41.

BROADHURST, P. L. Emotionality and the Yerkes-Dodson law. *J. exp. Psychol.*, 1957, **54**, 345–352.

CASTANEDA, A., & WORELL, L. Differential relation of latency and response vigor to stimulus similarity in brightness discrimination. *J. exp. Psychol.*, 1961, **61**, 309–314.

DUNCAN, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, **11**, 1–42.

ENDS, E. J., & PAGE, C. W. A study of functional relationships among measures of anxiety, ego strength and adjustment. *J. clin. Psychol.*, 1957, **13**, 148–150.

KORMAN, M. Ego strength and conflict discrimination: An experimental construct validation of the ego strength scale. *J. consult. Psychol.*, 1960, **24**, 294–298.

SCODEL, A., RATOOSH, P., & MINAS, J. S. Some personality correlates of decision making under conditions of risk. *Behav. Sci.*, 1959, **4**, 19–28.

WILLIAMS, H. L., & LAWRENCE, J. F. Comparison of the Rorschach and MMPI by means of factor analysis. *J. consult. Psychol.*, 1954, **18**, 193–197.

WORELL, L., & CASTANEDA, A. Individual differences in conflict resolution and the ease of learning. *Amer. Psychologist*, 1961, **16**, 400. (Abstract) (a)

WORELL, L., & CASTANEDA, A. Response to conflict as a function of response-defined anxiety. *J. Pers.*, 1961, **29**, 10–30. (b)

WORELL, L., & WORELL, JUDITH. An apparent source of drive in the methodology of motivational studies of reaction time and generalization. *J. gen. Psychol.*, in press.

(Received August 9, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 317–322

PERSONALITY AND HYPNOTIC SUSCEPTIBILITY[1]

PETER J. LANG AND A. DAVID LAZOVIK

University of Pittsburgh

An important facet of the researchers' interest in hypnosis concerns the relationship between personality and hypnotic susceptibility. A variety of measures have been used to evaluate the latter variable. They range from a single index such as postural sway (Eysenck, 1944) to multiple instruments such as those of Davis and Husband (1931), Barry, Mackinnon, and Murray (1931), and Friedlander and Sarbin (1938). More recently, Weitzenhoffer and Hilgard (1959) developed the Stanford Hypnotic Susceptibility Scale (SHS), which appears to have many advantages over former measures. The SHS is available in two forms. Each consists of 12 items, categorized into three groups: challenge, nonchallenge, and criterion items. The SHS can be administered in approximately 45 minutes and involves the actual induction of hypnosis. A statistical analysis of reliability, validity, and internal consistency is available on the original standardization group (Hilgard, Weitzenhoffer, Landes, & Moore, 1961).

Two main hypotheses have been raised concerning the relationship between personality and hypnosis. Eysenck (1960) described one of these as the linking of hysteria (and consequently extroversion) to primary suggestibility or hypnosis, and a second, asserting a connection between suggestibility and neuroticism. Eysenck held that the data supported the latter affirmation but concluded that suggestibility "fails to differentiate between extroverts and introverts" (Eysenck, 1960, p. 91). Subsequent research has not been wholly consistent with this analysis. Eysenck's conclusion that neurotics are more suggestible than normal subjects was derived mainly from studies in which postural sway was the dependent variable. More recent experiments employing this or similar measures, have reported little relationship between susceptibility and neuroticism when ataxic rate is controlled (Ingham, 1953).

Furthermore, the findings of Friedlander and Sarbin (1938) suggest a positive relationship between ease of inducing hypnosis and such traits as "amiability" and "sociability." Barber (1956) has advanced a "good guy" hypothesis of hypnotic susceptibility, based on high correlations between the latter variable and scores on scales of ascendance, sociability, emotional stability, and objectivity. Such traits have little in common with neuroticism, and some are positively related to extroversion. More recently Weitzenhoffer and Weitzenhoffer (1958) reported that they were unable to find any relationship at all between hypnotic susceptibility and a large battery of personality inventories.

The conflicting findings suggest the need for further research in this important area. The purposes of the present study are: (*a*) to investigate the relationship between hypnotic susceptibility (as defined by the SHS) and the personality variables to which it has been linked—extroversion, neuroticism, and "positive" personality traits; (*b*) to reexamine the internal consistency of the SHS; (*c*) to describe the distribution of hypnotic susceptibility in the current population; (*d*) to compare these latter findings with those of the original SHS standardization study.

PROCEDURE

A total of 46 college student volunteers (22 males and 24 females), attending introductory psychology, participated in this study. Their ages range from 18 to 28, with a mean of 20.61, $\sigma = 2.77$.

All subjects were administered Form A of the Stanford Hypnotic Scale by one of the two authors. Each experimenter hypnotized half of the total population. The two groups so defined were not significantly different in SHS scores and were treated as a unit in all analyses.[2]

[1] This study was carried out with the support of NIMH Grant M-3880.

[2] The major part of the statistical analysis was accomplished on the IBM 7070 computer at the Computation and Data Processing Center of the University of Pittsburgh. All correlations between

TABLE 1

PERCENTAGE OF SUBJECTS WHO SUCCESSFULLY COMPLETED EACH ITEM OF THE SHS AND CORRELATIONS BETWEEN INDIVIDUAL ITEMS AND THE TOTAL SCALE

Item	Percentage passing		Correlation with Total SHS[a]	
	L & L	W & H	L & L	W & H
1	78	69	.82	.38
2	48	58	.46	.57
3	85	81	.33	.63
4	17	14	.40	.75
5	35	32	.91	.72
6	41	32	.91	.83
7	74	70	.46	.51
8	30	23	.84	.79
9	52	35	.54	.55
10	33	30	.93	.79
11	48	49	.46	.60
12	43	32	.58	.69

Note.—L & L refers to the population evaluated in the current study; W & H refers to the findings of Weitzenhoffer and Hilgard (1959). In the latter case the correlations and percentages are based on both Forms A and B of the SHS. Only Form A was employed in this experiment.

[a] The correlations are all r_{bis}.

Thirty-two of these subjects were administered the following inventories, designed to evaluate the three dimensions of personality under consideration: (*a*) Bendig's Social Extroversion Scale,[3] the Pittsburgh revision of the Maudsley Extroversion scale (Bendig, 1959), the MMPI Social Introversion scale; (*b*) the Manifest Anxiety scale (Taylor, 1953), the Maudsley Neuroticism scale; (*c*) the Deference and Affiliation subscales of the Edwards Personal Preference Schedule (1953).[4] A high positive relationship between the SHS and the first two inventories, and a negative relationship with the third, would support the hypothesized connection between extroversion and hypnosis. On the other hand, a positive relationship between the SHS and the *MA* scale and Neuroticism scales would suggest a link between hypnosis and psychoneurotic behavior. Finally, significant positive correlation between the SHS and the Deference and Affiliation subscales of the Edwards schedule would indicate, for a college population, a relationship between positive social traits and hypnosis.

RESULTS

There is a high degree of consistency between the results of the present study and the standardization data of the SHS as reported by Hilgard et al. (1959,[5] 1961). These latter authors reported no sex differences in SHS scores ($t = 0.35$) and did not treat males and females separately. This result was confirmed in the current sample ($t = 0.74$) and sexes were combined in subsequent analyses of the SHS. As can be seen in Figure 1, the bimodal distribution of susceptibility scores reported by Hilgard and his co-workers was also obtained in the present experiment. Inspection of Table 1 reveals considerable similarity between the two studies in the percentage of subjects passing individual items. A rho of .78 is obtained when the rank differences are analyzed. Table 1 also permits a comparison

[5] E. R. Hilgard, A. M. Weitzenhoffer, J. Landes, and Rosemarie K. Moore, personal communication, 1959.

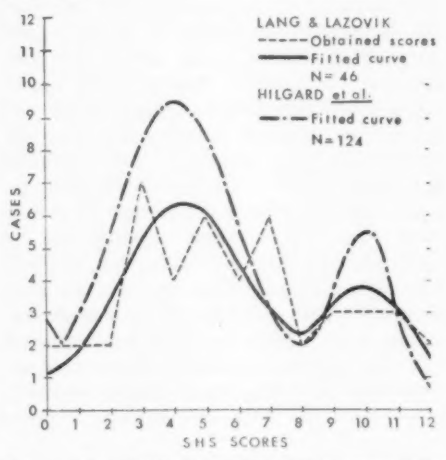

FIG. 1. The distribution of hypnotic susceptibility in the college population.

continuous variables (Total SHS and personality scales) and discontinuous variables (SHS items) were originally stated as point biserial coefficients. These were subsequently converted to r_{bis} to allow for comparison with the statistics of the original SHS standardization group, except where otherwise indicated.

[3] This is an improved scale of 20 items derived, after repeated item analysis, from the 60 items which make up the MMPI Social Introversion scale. Scoring was reversed to give a measure of extroversion. Personal Communication from A. W. Bendig, 1961.

[4] In addition to these inventories nine additional scales of the Edwards schedule and two scales (restraint and thoughtfulness) developed by Guilford (1949) were available. The MMPI was also administered to all subjects prior to the SHS in order to screen out individuals who might respond to hypnosis in a psychotic manner. Six subjects were eliminated from the study who scored above 70 on either the *Pa* or *Sc* scales of the MMPI.

of the original standardization data and the present findings with respect to the correlations between individual items and total score. In only three cases, postural sway, arm immobilization, and hand lowering, are the results appreciably different. In Table 2 items are classified according to criterial, nonchallenge, and challenge categories. The latter grouping shows the highest intercorrelation among items and the highest correlation with total score in both studies.

In addition to the correlation between the total SHS and the personality battery used in this experiment, estimates of the relationship between the battery and the individual items of the SHS were obtained. Of particular interest is Item 1, postural sway, which has frequently been used as a single index of hypnotic susceptibility in previous research. Table 3 presents the correlations between the Total SHS, the postural sway item, and tests pertaining to the three main hypotheses under examination. Because scores on personality tests are frequently sex-related, correlations obtained for the males, females, and total sample are reported separately.

The present results reveal a negative relationship between the indices of neuroticism and both postural sway and the Total SHS. They also indicate a low positive relationship between extroversion and these measures of hypnotic susceptibility. The subscales reported from the Edwards schedule are the two yielding the highest correlations with the Total SHS. Neither scale of the Guilford nor the remainder of the Edwards subscales yielded correlation coefficients which approached significance.

TABLE 2

INTERCORRELATION AMONG ITEMS WITHIN CHALLENGE, NONCHALLENGE, AND CRITERIAL CATEGORIES AND THE CORRELATION OF THE CATEGORIES TO TOTAL SCORE

Category	Items Intercorrelation [a]		Correlation with Total [b]	
	L & L	H et al.	L & L	H et al.
Challenge (Items 4, 5, 6, 8 10)	.94	.83	.85	.84
Nonchallenge (Items 1, 2, 3, 7)	.52	.57	.55	.56
Criterial (Items 9, 11, 12)	.28	.57	.53	.66

[a] Hilgard (Hilgard, Weitzenhoffer, Landes, & Moore, personal communication, 1959) present the arithmetic average of z scores converted from the original tetrachoric interitem correlations. In the present experiment the IBM 7070 produced phi coefficients to describe the relationships between items. These were subsequently converted to an estimate of Pearson's r (Edwards, 1946), which closely approximates the tetrachoric correlations obtained by Hilgard. The means of the z scores were then obtained to provide the statistics reported in this table.

[b] In both studies these statistics are the arithmetic means of z scores derived from biserial correlations.

TABLE 3

CORRELATIONS BETWEEN PERSONALITY DIMENSIONS AND THE STANFORD HYPNOTIC SUSCEPTIBILITY SCALE

Personality dimension	SHS			Postural sway item (SHS)		
	Males	Females	Total	Males	Females	Total
Extroversion						
Social Extroversion (Bendig)	.18	.50*	.35*	.47	.32	.35*
Social Introversion (MMPI)	−.13	−.51*	−.30	−.41	−.22	−.33
Maudsley inventory	.16	.20	.21	.24	.19	.25
Neuroticism						
Maudsley inventory	−.20	−.38	−.24	−.55*	−.09	−.37*
Manifest Anxiety scale	−.27	−.39	−.29	−.50*	.05	−.32
Edwards schedule						
Deference	.48	.15	.27	.50*	.16	.33
Affiliation	.41	.10	.36*	.50*	−.26	.34

Note.—Correlations involving the Total SHS are Pearson r's; $r_{p\text{-}bis}$ is reported for the postural sway item. These statistics are based on data obtained from 16 males, 16 females forming a total N of 32.

* Significant at .05 level.

DISCUSSION

The findings of the present experiment, relevant to the structure of the SHS, are consistent with those of the original standardization sample. This is true of the distribution of susceptibility scores, the difficulty of individual items, the general form of the intercorrelation matrix, and the majority of the correlations between individual items and the total scale. Three of the latter correlations did vary from those reported by Weitzenhoffer and Hilgard (1959). However, it is interesting to note that they all involved items which appear early in the scale. Postural sway is administered before actual hypnosis is induced, and hand lowering and arm immobilization are the first two items presented after eye closure. The instability of these items may reflect the subject's unease in the situation or initial apprehension about hypnosis. This factor probably dissipates with successive items, as the subject becomes more involved in the hypnotic experience. Initial unease would have had less effect on the correlations reported in the original study. The statistics cited by Weitzenhoffer and Hilgard (1959) are actually based on two forms of each item, one of which was administered at a second session. Presumably, rapport problems would have been considerably reduced at this time.

The personality inventory correlations indicate a significant positive relationship between extroversion, traits suggesting an interest in and consideration for others, and hypnotic susceptibility. Sex differences are also implied. However, when the population is subdivided by sex the number of subjects per group ($N = 16$) becomes so small that differences may be attributable to factors idiosyncratic to the sample. For example, the generally low correlations between the personality inventories and the female postural sway score are clearly a function of the fact that so few girls failed this item. The range of scores was simply not great enough to provide the basis for an accurate correlational estimate.

Neuroticism was found to bear a slight negative relationship to hypnotic susceptibility, as measured by both the postural sway item and by the Total SHS. The failure to find a positive relationship is consistent with a host of studies reviewed by Weitzenhoffer (1953) and Heilizer (1960). It is in contradiction with the work of Eysenck (1960) and studies in which actual neurotic populations are employed (Himmelweit, Desai, & Petris, 1946). In commenting on the inconsistency, Weitzenhoffer points out that a scale measure is used in one type of study, and in the other the neurotic is defined by psychiatric diagnosis. He suggests that "a tendency to neuroticism," measured by a scale in the normal population, may be different from neuroticism itself, palpable to psychiatric diagnosis. On the other hand, Heilizer (1960) holds that the conflict arises from the ubiquitous failure to control drug usage in studies of pathological groups.

An explanation not previously considered involves the distinction between passive and active hypnotic states, proposed by early workers in hypnosis (Bernheim, 1880; Binet & Féré, 1888). The active trance state is characterized by the subject's eager participation in hypnosis, absence of delay in responding to suggestions, and a clear indication that effort is being expended attempting to overcome inhibitory suggestions. Conversely, the subject who achieves a passive trance expends little effort and shows delay in responding to suggestions. He appears sluggish, concerned with repose, and inclined to resist disturbance. More recent studies of these two hypnotic states are reported by Davis and Kantor (1935) and Eysenck and Furneaux (1945).

In a 1937 article, R. W. White reported a relationship between these two states and certain dimensions of personality. According to White, the active hypnotic subject is judged high in the traits of "deference" and "affiliation" (i.e., he is willing to go along with authority; he seeks out others, and enjoys participating in social situations). In contrast, the passive hypnotic state is associated with such traits as "abasement" (willingness to accept defeat, one's own inferiority), infavoidance (the tendency to retreat from difficulty) and anxiety.

It is interesting to note that the traits of affiliation and deference are the only two from either the Edwards or Guilford schedule that show a positive relationship to hypnotic susceptibility in the present study. The three

traits White connects to the passive trance state are all involved in estimates of neuroticism, a trait negatively correlated with hypnotizability in this sample.

In discussing their study, Weitzenhoffer and Weitzenhoffer (1958) suggested that the failure to find personality correlates of hypnotic susceptibility was attributable to the fact that passive and active hypnotic subjects were not separately analyzed. They further proposed that studies employing small samples might arbitrarily select out the active type. A review of notes taken during the hypnotic sessions of all 46 subjects who participated in this research, yielded only two individuals who gave clear evidence of the passive trance state. On the other hand, almost a quarter of the total population could be classified as having experienced active trance states. These findings suggest that a loading of active hypnotic subjects in the present population may account for the particular relationship obtained between hypnotizability and personality.

These facts also lead to the hypothesis that the difference in correlates of hypnotizability found by previous studies may be attributable to the different proportions of active and passive hypnotic subjects in the true neurotic and the normal population. One might anticipate that the behavior of a good hypnotic subject drawn from a sample of college volunteers is more likely to be characterized by an active trance state; [6] the highly susceptible neurotic would be of the passive type. Thus within the normal population susceptibility would appear to be highly correlated with extroversion and positive social traits. In the pathological group it would be related to increasing neuroticism and withdrawal. The difference in these findings can be attributed to the fact that two different hypnotic states are being measured.

Recently Furneaux and Gibson (1961) and Furneaux (1961) have reported data consistent with this interpretation. The latter author states that, "In the group studies, the body-sway scores of stable extraverts and neurotic introverts tended to be large, whereas they were smaller for stable introverts and neurotic extraverts" (p. 213). It is reasonable to assume that body-sway correlates highly with both types of hypnotic state (active and passive). If so, this is exactly the relationship between extroversion, neuroticisim, and hypnotic susceptibility our hypothesis anticipates in a large population containing both active and passive hypnotic subjects.

However, Furneaux (1961) suggests that this finding can be best explained in terms of differences in drive level, associated with particular combinations of extroversion and neuroticism. The relationship between drive factors and the active-passive dichotomy should be explored. Furthermore, an experiment which employs both normal and diagnosed neurotic populations, controls the drug variable and categorizes hypnotic behavior along the active-passive dimension, would be of considerable help in clarifying the findings in this area.

[6] Hilgard et al. (1961) report little evidence for the passive-active dichotomy. However, college volunteers were employed in their work, and experience with this sample suggests that passive subjects are rare in this population.

SUMMARY

The investigation of the internal structure of the Stanford Hypnotic Scale (SHS) and the distribution of susceptibility in a college population, yielded results consistent with those of the original standardization study. It was also found that hypnotic susceptibility was negatively related to neuroticism and positively related to measures of extroversion, deference and affiliation. An explanation was proposed of conflicting findings concerning the relationship between hypnotic susceptibility and personality.

REFERENCES

BARBER, T. X. A note on hypnotizability and personality traits. *J. clin. exp. Hypnosis.*, 1956, **4**, 109–114.

BARRY, H., JR., MCKINNON, D. W., & MURRAY, H. A., JR. Studies in typological relations. *Hum. Biol.*, 1931, **13**, 1–36.

BENDIG, A. W. College norms for the concurrent validity of the Pittsburgh revisions of the Maudsley Personality Inventory. *J. psychol. Stud.*, 1959, **11**, 12–17.

BERNHEIM, H. *Suggestive therapeutics* (Rev. ed.) New York: Putnam, 1880. (Trans. by C. A. Herter)

BINET, A., & FÉRÉ, C. *Animal magnetism*. New York: Appleton, 1888.

DAVIS, L. W., & HUSBAND, R. W. A study of hypnotic susceptibility in relation to personality traits. *J. abnorm. soc. Psychol.*, 1931, **26**, 175–182.

DAVIS, R. C., & KANTOR, J. R. Skin resistance during hypnotic states. *J. gen. Psychol.*, 1935, **13**, 62–81.

EDWARDS, A. L. *Statistical analysis for students in psychology and education*. New York: Rinehart, 1946.

EDWARDS, A. L. *Manual for the personal preference schedule*. New York: Psychological Corporation, 1953.

EYSENCK, H. J. States of high suggestibility and the neuroses. *Amer. J. Psychol.*, 1944, **57**, 406–411.

EYSENCK, H. J. *The structure of human personality*. (2nd ed.) New York: Wiley, 1960.

EYSENCK, H. J., & FURNEAUX, W. D. Primary and secondary suggestibility: An experimental and statistical study. *J. exp. Psychol.*, 1945, **35**, 485–503.

FRIEDLANDER, J. W., & SARBIN, R. T. The depth of hypnosis. *J. abnorm. soc. Psychol.*, 1938, **33**, 281–294.

FURNEAUX, W. D. Neuroticism, extraversion, drive, and suggestibility. *Int. J. clin. exp. Hypnosis*, 1961, **9**, 195–214.

FURNEAUX, W. D., & GIBSON, H. B. The Maudsley Personality Inventory as a predictor of susceptibility to hypnosis. *Int. J. clin. exp. Hypnosis*, 1961, **9**, 167–177.

GUILFORD, J. P., & ZIMMERMAN, W. S. *The Guilford-Zimmerman Temperament Survey*. Beverly Hills, Calif.: Sheridan Supply, 1949.

HEILIZER, F. An exploration of the relationship between hypnotizability and anxiety and/or neuroticism. *J. consult. Psychol.*, 1960, **24**, 432–436.

HILGARD, E. R., WEITZENHOFFER, A. M., LANDES, J., & MOORE, ROSEMARIE K. The distribution of susceptibility to hypnosis in a student population: A study using the Stanford hypnotic susceptibility scale. *Psychol. Monogr.*, 1961, **75**(8, Whole No. 512).

HIMMELWEIT, H. T., DESAI, M., & PETRIS, A. An experimental investigation of neuroticism. *J. Pers.*, 1946, **15**, 173–196.

INGHAM, J. G. Psychoneurosis and suggestibility. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

TAYLOR, JANET A. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, **48**, 285–290.

WEITZENHOFFER, A. M. *Hypnotism*. New York: Wiley, 1953.

WEITZENHOFFER, A. M., & HILGARD, E. R. *Stanford hypnotic susceptibility scale*. Palo Alto, Calif.: Consulting Psychologists, 1959.

WEITZENHOFFER, A. M., & WEITZENHOFFER, G. B. Personality and hypnotic susceptibility. *Amer. J. clin. Hypnosis*, 1958, **1**, 79–82.

WHITE, R. W. Two types of hypnotic trance and their personality correlates. *J. Psychol.*, 1937, **3**, 279–289.

(Received August 23, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 323-330

THE INFLUENCE OF SYMBOLIC STIMULUS VALUE ON PERCEIVED SIZE IN CHRONIC SCHIZOPHRENIA[1]

VLADIMIR PISHKIN, THOMAS E. SMITH[2]

Research Laboratory, VAH, Tomah, Wisconsin

AND H. W. LEIBOWITZ[3]

University of Wisconsin

It is an accepted fact that the impairment of chronic schizophrenia is both diverse and complex. Deviation has been demonstrated in the areas of cognition, perception, motivation, and emotion. Some theorists have argued for the primacy of deficit in one of these areas, and consequent impairment of the others. Psychoanalytic theory considers schizophrenia as an affective disorder (Eissler, 1954; Freud, 1947). Other theorists have argued for the primacy of deficit in cognitive function (Arieti, 1955; Goldstein, 1943); while still others theorize that the most basic deficit in schizophrenia is motivational (Cohen, 1956; Huston & Shakow, 1946).

Probably the most recent aspect of human responsivity which has been considered as the primary impairment in schizophrenia is perception (Klein, 1951; Lovinger, 1956). Cooper (1960) states this position quite concisely when she writes, "While it may be true that schizophrenic impairment involves cognitive difficulties, or emotional conflicts, or both, another possible hypothesis is that schizophrenia involves, more simply and directly, impaired perceptual functioning" (p. 209).

The major problem in investigation of the above hypothesis is that of minimizing the effects of any cognitive or emotional deficit. Most investigators have tried to reduce the effects of cognitive, affective, and motivational impairment by using very simple, objective problems. Studies have been reported using the autokinetic effect (Voth, 1947), weight judgment (Salzinger, 1957), and size judgment (Cooper, 1960).

The purpose of the present experiment is to extend previous investigations of size judgments in chronic schizophrenics. The results of studies comparing normals and schizophrenics with regard to this classical phenomenon of perception are somewhat contradictory (Rabin & King, 1958). Sanders and Pacht (1952) reported "overconstancy" in schizophrenia, while Weckowicz (1957) reported the opposite. Cooper (1960) recently reported a decrease in accuracy and an increase in variability on simple size comparison tasks. These results are in disagreement with a recent study by Leibowitz and Pishkin (1961), in which no differences were found between normals and schizophrenics on a simple task of judging sizes of wooden dowels under well lighted conditions. The conclusions reached on the basis of that study were, "whatever disorders of perception, thinking, or behavior they (chronic schizophrenics) suffer from, their size constancy is unaffected" (p. 198).

At least two explanations of the contradictory results of these studies are possible. First, there may be differences in the subject populations, especially the patient groups, and there is a possibility of different syndromes within the broad category of schizophrenia. For example, Sanders and Pacht (1952) used outpatients, while Cooper (1960) used a schizophrenic population which was predominantly composed of patients of less than 2 years' hospitalization. Neither of these groups is comparable to the population used by Leibowitz and Pishkin (1961) which had an

[1] This project was supported by the Veterans Administration Research Program (8200 Funds) and Grant M-1090 from the United States Public Health Service, National Institutes of Health, National Institute of Mental Health.

The authors are indebted to Lewis M. Jones for his assistance in processing of the data.

[2] Now at Central Michigan University.

[3] Now with International Business Machines Corporation, Federal Systems Division; Bethesda, Maryland.

average length of hospitalization of close to 9 years.

The second reason for the contradictory results in previous studies may be due to differences among the stimulus objects employed. It is quite possible that the failure to find deficit in schizophrenics is due to the use of neutral stimuli. Raush's hypothesis that affective aspects have a greater influence on size judgment of schizophrenics than normals was supported in his studies (Raush, 1952, 1956) requiring estimates of both neutral objects and objects having a more affective value. Zahn (1959) has also stressed the utility of size judgment technique in evaluation of affective variables' influences on schizophrenics' behavior. In the earlier study (Leibowitz & Pishkin, 1961) the experimental stimuli were simple wooden dowels, which presumably have no symbolic or emotional connotation for the patients. It is possible that when stimuli become emotionally charged the result is a deficit in the performance of schizophrenics. However, one could also expect an increase in differences between schizophrenics and normals in size matching by minimizing cue conditions. Whitman (1954) has reported a greater impairment in the concept formation performance of schizophrenics when stimuli are of an interpersonal nature. If the above argument is valid, then it follows that schizophrenic impairment is not primarily a deficit in purely perceptual functioning, but is more likely the result of a complex interaction between perceptual, cognitive, emotional, and motivational processes. In fact, it may be that the primary impairment is in one of the other areas of function, and that the deficit in perception is secondary. Weckowicz and Blewett (1959) have emphasized the aspect of cue selection as an influencing factor on perceptual performance of certain groups of schizophrenics. In the present experiment, short flash durations were employed in illuminating the visual field. In this way the subjective aspects of size judgment on cue selection were maximized.

This study was designed to investigate the possibility that an increase in the affective value of stimuli may result in a decrease in accuracy of size judgments of chronic schizophrenics as compared to normals. The technique involved required the subject to indicate whether a human shaped plywood silhouette is larger, smaller, or the same size as a three-dimensional clothed dummy figure. The patient population was essentially the same as that of the previous study (Leibowitz & Pishkin, 1961). They were all chronic schizophrenics who had demonstrated marked withdrawal symptoms. It was assumed that the present task involving human-like figures would be more affect arousing than the earlier task using dowel rods. In addition, it was expected that differences between schizophrenics and normals would be maximized with the decrease in visual cues.

METHOD

Subjects

The experimental group was composed of 40 male psychiatric patients of the Veterans Administration Hospital in Tomah, Wisconsin, carrying a diagnosis of chronic undifferentiated schizophrenia. Thirty-five of these patients had served as subjects in the previous study (Leibowitz & Pishkin, 1961).[4] None of the patients had a history of brain damage, organic pathology, or visual defect. None were receiving EST, but all of them were on some drug treatment program. They were selected by the psychiatric ward team to be mainly characterized by withdrawal. The mean age of the 40 patients was 40.87, with a range from 24 years to 56 years. The average length of hospitalization was 9.7 years, with a range from 3 years to 14 years.

The control group was composed of 40 male psychiatric nursing aides, selected at random from the hospital staff. The mean age of the 40 aides was 38.88, with a range from 22 years to 58 years.

Experimental Conditions

The experiment was conducted in a hospital staff room with dimensions of 48 feet × 28 feet. All subjects were seen individually, at night, and shades were drawn to completely darken the room during the actual size judgments. The standard stimulus was a 65 inch three dimensional figure, mounted on a 2 inch platform, bringing the total height of the figure to 67 inches. This figure was very lifelike; the colors were natural and it was used for the purpose of training student nurses.[5] In order to vary the sym-

[4] The two groups of patients who were exposed to either "aide" ($N = 17$) or "patient" ($N = 18$) stimulus in this study were equated on the previous studies' dowel task. The combined mean deviation scores from the standard dowel stimulus were .82 and .79 for the aide and the patient groups, respectively.

[5] It is manufactured by the Chase Manufacturing Company for training in medical care of patients.

bolic value, this standard was dressed in a white uniform identical to those customarily worn by the male nursing aides in the hospital for one half of the subjects from each group, and was dressed in the green uniform issued VA patients for the remaining half of the subjects from each group. The test stimuli were five silhouettes, cut from ¼ inch plywood and painted battleship gray. The silhouettes were cut proportionately to the frontal view of the dummy, and were of varying size. The size varied in 2 inch decrements, so that the five figures were of 63, 65, 67, 69, and 71 inches in height, and were proportional to the size of the dummy.

The standard stimulus was 12 feet to the left of the subject's position, and the comparison stimuli were 40 feet directly in front of this position. The background for both the standard and the experimental figures was a white screen. The subjects were ushered into the room while the overhead lights were on, and were seated in a chair facing the screen in front of which the test stimuli were to appear. They were told that a series of figures would be presented, the same shape as the dummy at their left, but of differing size. They were to indicate verbally whether the silhouette was larger, smaller, or the same size as the dummy figure. The subjects were told that the silhouette would be placed in front of the screen in front of them, right after the lights were turned out, and when they were told to start, they could illuminate both the dummy figure and the silhouette by depressing a lever of the photostimulator which was before them. This was a standard Grass photostimulator, which produced a flash of 1,500,000 candlepower (at a distance of 10 feet) for a duration of 10 microseconds. This short duration was chosen in order to maximize subjective aspects of size judgment. It was placed so that the beam directly bisected the 90 degree angle between the dummy figure and the test silhouettes. The subjects were instructed that they could flash the light as often as they desired, until they were ready to make the decision as to whether the silhouette was larger, smaller, or the same size as the standard figure. They were informed that the figure to be judged would not change until they had made their decision. Then the overhead lights would go on, and shortly after they went off again, the test stimuli would change and they would be told to illuminate the field again.

Each subject was presented with all five silhouettes. The order of presentation was determined by a 5×5 Latin square pattern, with four subjects from the patient group and four subjects from the normal group going through each order when the standard figure was dressed as an aide, and four subjects from the patient group and four subjects from the normal group going through each order when the standard figure was dressed as a patient. There was no difficulty encountered in testing the subjects. They all cooperated well, and did not demonstrate any difficulty in understanding the instructions. There were no signs of complaint or fatigue during the experiment, which took about 5 minutes per subject.

RESULTS

Each subject's performance on the experimental task was scored along the three dimensions. These were: size judgment, number of errors, and frequency of illuminations. The size judgment score was the average height of the silhouette figure which appeared to a subject to be the same size as the standard figure; the error score was the number of wrong judgments on the five trials, and the illumination score was the number of times the subject illuminated the field by means of the photostimulator before making a size judgment on each of the five trials.

Variance ratio (McNemar, 1955) was used in order to estimate the homogeneity of variances among the four treatment groups in each of the three measures, i.e., size, errors, and illuminations. The variances were homogeneous among the groups for both the size and the error scores. However, there was significant heterogeneity of variance between two cells of illumination scores (schizophrenics with aide stimulus and normals with aide sti ulus). These variances were compared, producing significant heterogeneity ($F = 20.17$, 19 and 19 df, $p < .001$). The variances were stabilized by square root transformation (Edwards, 1958) of the four illumination scores, producing sufficient reduction in heterogeneity. Table 1 represents means and standard deviations of the three measures.

The mean size judgment scores for the four experimental groupings are shown in Figure 1. Table 2 presents a summary of the analysis of

TABLE 1

SIZE, ERROR, AND ILLUMINATION SCORES OF SCHIZOPHRENIC AND NORMAL SUBJECTS FOR THE TWO STIMULUS CONDITIONS

	Size		Error		Illumination [a]	
Stimulus	*M*	*SD*	*M*	*SD*	*M*	*SD*
			Schizophrenics			
Patient	66.45	2.20	3.05	1.31	4.26	2.01
Aide	68.65	2.52	2.41	1.06	7.46	2.73
			Normals			
Patient	67.20	2.51	2.40	.90	3.88	1.58
Aide	67.75	1.80	2.50	.96	3.82	1.62

[a] Square root transformation was applied to the raw Illumination scores.

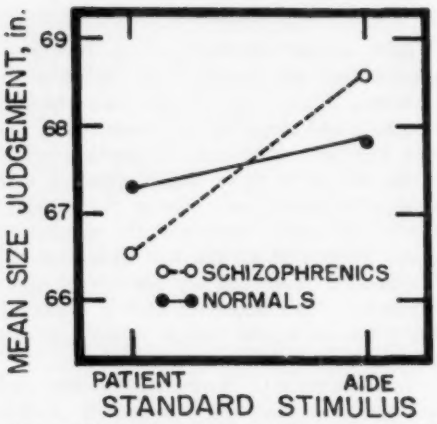

FIG. 1. Mean size judgments of standard stimuli by schizophrenics and normals. The size of the two standard stimuli was 67 inches.

variance of these size judgment scores. Most important to the hypothesis under consideration is the finding that there was no significant difference between the schizophrenics and the normals. Even though both groups of subjects were close in their size judgment scores by producing average numerical values of about 67, it was important to determine when the schizophrenics were more diffuse in arriving at their mean size score, i.e., they could have responded to all five comparison stimuli as "same" and thus would have come up with an accurate score. Table 3 represents the frequencies of "smaller," "same," and "larger" responses to the five stimuli by the two populations. Inspection of this table reveals that the average numerical values to which the subjects responded "same" were not influenced by differences in degree to which the two subject groups may have been differentially diffuse in their judgment. However,

TABLE 2

SUMMARY OF THE ANALYSIS OF VARIANCE OF SIZE JUDGMENT SCORES

Source of variation	*df*	*MS*	*F*
Between groups	1	.12	.03
Between conditions	1	37.82	7.67*
Interaction	1	13.61	2.76
Within	76	4.93	

* $p < .05$.

there was a significant difference in the size judgments under the two experimental conditions ($p < .01$). Both schizophrenics and normals tended to judge the larger test silhouettes equal in size to the standard figure clothed as an aide. This change in size judgment was greater for the schizophrenics. In fact, it can be seen in Figure 1 that when the standard was clothed as a patient the schizophrenics judged a smaller test silhouette of equal size than the normals, and when the standard was clothed as an aide the schizophrenics judged a larger test silhouette of equal size than the normals.

Figure 2 shows the mean error score for the four experimental groupings, and Table 4 presents a summary of the analysis of variance of these error scores. Overall, schizophrenics made significantly more errors than normals ($p < .05$). However, it can be seen in Figure 2, and in the significant interaction ($p < .05$), that the real difference was in terms of the increase in errors by the schizophrenics when comparing the test figures to the standard clothed as a patient. In fact, schizophrenics made fewer errors than aides when the test standard was clothed as an aide.

The mean illumination scores for the four groupings are shown in Figure 3. The sum-

TABLE 3

FREQUENCY DISTRIBUTION OF RESPONSES TO THE FIVE COMPARISON STIMULI FOR SCHIZOPHRENICS AND NORMALS

Stimulus size	63″			65″			67″			69″			71″		
Judgment	−	0	+	−	0	+	−	0	+	−	0	+	−	0	+
Schizophrenics	31	7	2	21	12	7	15	17	8	9	14	17	9	10	21
Normals	29	10	1	22	13	5	18	18	4	8	13	19	11	11	18

Note.—The signs in the table represent the following responses: smaller (−), same (O), and larger (+).

TABLE 4

SUMMARY OF THE ANALYSIS OF VARIANCE OF ERROR SCORES

Source of variation	*df*	*MS*	*F*
Between groups	1	4.05	4.45*
Between conditions	1	1.25	1.37
Interaction	1	4.05	4.45*
Within	76	.91	

* $p < .05$.

mary table of the analysis of variance on the transformed illumination scores is shown in Table 5, and again it can be seen that there was a significant difference between schizophrenics and normals ($p < .001$). The schizophrenics illuminated the field more frequently under both of the experimental conditions, but the greatest difference was when the standard stimulus was clothed as an aide. Also, the main effect of methods was significant ($p < .001$), due to greater frequency of illuminations with aide as a stimulus. It was apparent that the significant interaction ($p < .001$) between the Groups and the Conditions was responsible for the significance of the Groups main effect.

Pearson *r*'s were computed between size judgment scores, errors, and frequencies of illuminations within each of the four treatment groups (i.e., schizophrenics with patient stimulus, schizophrenics with aide stimulus,

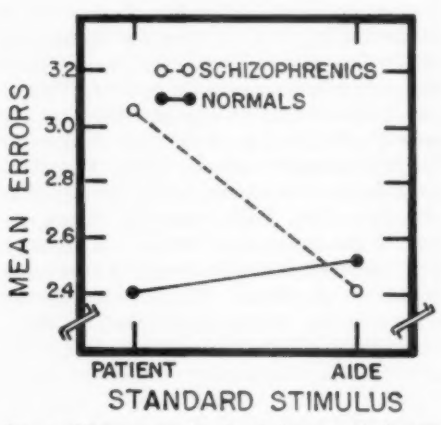

FIG. 2. Mean errors in size judgment of the two standard stimuli by schizophrenics and normals.

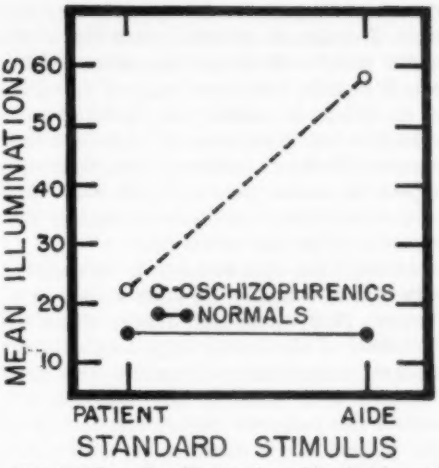

FIG. 3. Mean illuminations by schizophrenics and normals in making size judgment of the two standard stimuli.

normals with patient stimulus, and normals with aide stimulus). The significant correlations were with schizophrenic subjects between errors and illuminations with aide stimulus ($r = .48$, $p < .05$) and with normals between errors and illuminations, also with aide stimulus ($r = .52$, $p < .05$). The other 10 correlations were close to zero. Pearson *r*'s were also computed between the three scores combining both subject populations. There were no significant correlations. The *r*'s were as follows: illuminations and errors, $r = .02$; illuminations and size, $r = -.09$; size and errors, $r = .04$.

DISCUSSION

The results of this study indicate no significant differences in the size judgments of

TABLE 5

SUMMARY OF THE ANALYSIS OF VARIANCE OF ILLUMINATION SCORES

Source of variation	*df*	*MS*	*F*
Between groups	1	79.99	21.27**
Between conditions	1	50.24	13.36**
Interaction	1	52.16	13.87**
Within	76	3.76	

Note.—Square root transformation was applied to raw scores.
** $p < .001$.

chronic schizophrenics in comparison to normals. The data is consistent with that of the earlier study with similar population (Leibowitz & Pishkin, 1961), and suggests that there is no deficit in making size judgments involved in the impairment of chronic schizophrenia. However, patient-control differences appear to evolve from schizophrenics' difficulty in selecting cues under minimized illumination of the perceptual field.

Although the data seems to be inconsistent with that of a number of other studies (e.g., Cooper, 1960; Lovinger, 1956), closer examination of the results helps to clarify the apparent inconsistencies. Data from this study indicated that although the schizophrenics made a size judgment very similar to that of the normals, they made more errors, and found it necessary to illuminate the field more frequently, before arriving at that judgment. Equally important, schizophrenics made more frequent illuminations to the aide stimulus than normals. One possible interpretation might be that schizophrenics may have reacted to the aide stimulus with more "vigilance" and with apathy to the patient stimulus. Since only one estimate of size for each of the five comparison stimuli was made after a number of illuminations by each individual tested the error score was not confounded by the frequency of illumination. Significant, positive correlations between error and illumination scores for both populations with aide as a stimulus would indicate that the inaccuracy in making judgments was accompanied by the necessity to illuminate the field more frequently in order to receive sufficient cues before making a decision. These positive correlations between errors and illuminations support the notion that the subjects received less information concerning dummy size with each illumination of the field, and therefore, the increase of illumination frequency was accompanied with increase in errors. It is quite plausible that the subjects were misinformed more often under aide stimulus condition in that their selection of cues was not valid, thus producing more errors. Therefore, it is implied that isolation of relevant cues in this condition was more difficult. This view regarding cue validity and misinformation was supported earlier (Pishkin, 1960; Pishkin, 1961). There is a possible flaw in the present methodology. Since both stimulus dummies were revealed against a white screen, the figure-ground variable may have had an influence on the subject's judgment. In the case of aide stimulus, there was a possibility of discontinuity with the background. On the other hand, a patient dummy was dressed in a green uniform. This possibility cannot be ruled out, and it would support the importance of cue selection. Nevertheless, the other alternative would have been to have matching figure-ground fields, in which case the background for the two stimulus conditions would not have been controlled.

Cooper (1960) compared the perceptual judgments of schizophrenics and normals, and reported significant differences in accuracy and variability. As in the present study, she found that schizophrenics make more errors. She speculated that the increase of errors is the result of perceptual impairment. An alternate explanation of these findings is that the schizophrenics have greater difficulty in attending to, and selecting out, the relevant cues in the visual field. This hypothesis is given some support by the present finding that the schizophrenics find it necessary to illuminate the test field more frequently before making a size judgment. It is possible that the schizophrenic has difficulty in selection of the relevant cues and rejection of the irrelevant cues during the period of illumination. Explanations for this difficulty might include impairment in attention, an increase in distractability, and an increase in internal stimulation. Consequently, the present findings cannot be interpreted in term of a "pure" perceptual deficit. It is also plausible that the potency of stimuli is subjectively weaker for the schizophrenic subjects. Thus, it is compensated by them through more frequent illuminations. This finding may be closely related to the notion that the channel capacity of schizophrenics deteriorates with increasing chronicity (Lydecker, Pishkin, & Martin, 1961). In an earlier paper, Pishkin, Armstrong-Ressy, Aller, and Comstock (1960) suggested that deterioration of schizophrenic's ability to attend to environmental cues increases with chronicity. It may well be that attention or the amount of information trans-

mitted in a single illumination of the field, used in this study, is related to duration of illness and not to schizophrenic pathology, per se. It may be safely assumed that the perceptual environment of a long-term schizophrenic patient becomes limited (mean duration of hospitalization for this subject population was 9.7 years), and progressively less demanding on the subject's decision-making capacities such as those required by the present task. It would be interesting to investigate the relationship between length of hospitalization and ability to make size judgments and simple discriminations by schizophrenic subjects.

Weckowicz and Blewett (1959) have reported significant correlations between size judgments, ability to form concepts, and span of attention in schizophrenics. It appears that the schizophrenic is less able to attend selectively, and consequently performs more poorly on many different kinds of tasks. The deficit in selecting and attending may be the result of distraction to irrelevant aspects of the external environment, or distraction due to internal symbolic arousal. In either case there could be a decrease in the ability to attend selectively, not only at the cognitive level, but at the sensory level (e.g., Hernandez-Peon, Sherrer, & Jouvet, 1956). This would result in less information input, and consequent poor performance.

The results also indicated that the schizophrenics made the most errors when they were comparing the test stimuli to the standard clothed as a patient. It is highly probable that the patient clothed as a standard had greater symbolic value, and aroused numerous distracting thoughts concerning self-evaluations. This supports the hypothesis of Bruner (1951) that an increasing concern for the self might lead to a breakdown in perceptual functioning. Furthermore, both groups saw the test figures as larger when comparing them to the standard dressed as an aide. Perhaps both the schizophrenics and the aides perceived the patient as smaller because of negative value judgments concerned with the importance, adequacy, and general worth of the mentally ill person. It would follow that such a judgment would be more disturbing to the schizophrenic, and consequently his performance would be poorer. Although there was no significant difference between the groups in their size judgment of both patient and aide standard stimuli, it is noteworthy that both the patient and the aide subjects judged the aide stimulus larger than he really was. Perhaps, the schizophrenic subjects' self-esteem was lower than that of the normal subjects.

In general, the results of the study support the hypothesis that it is the symbolic or emotional value of the stimulus which accounts for schizophrenic deficit on perceptual tasks. Basic perceptual functions, such as size constancy, and size judgments, are unaffected. This is probably because the required abilities are fully developed prior to the development of schizophrenic process, and are not affected by the social and emotional withdrawal accompanying the syndrome itself. This explanation was suggested earlier by Leibowitz and Pishkin (1961). However, the present findings suggest that patient-control differences lie in their abilities to select relevant cues.

SUMMARY

The ability to make size judgment comparisons as a function of the height of the test figure and the symbolic value of the standard figure was determined for a group of 40 chronic, undifferentiated schizophrenics, as well as for a control group of 40 psychiatric aides. The symbolic value was varied by means of using different sets of clothing (patient or aide) on the standard stimulus.

The highlights of the major findings were:

1. There were no significant differences in the size judgments made by the two groups.

2. The schizophrenics made significantly more errors, and found it necessary to illuminate the field more frequently, before arriving at the size judgment. It was suggested that the increase in errors, and the increase in the number of illuminations, for the schizophrenic group were a result of their distraction due to irrelevant aspects of the stimulus field, irrelevant internal processes, and generalized inability to make judgments. The schizophrenics could not obtain the same amount of information regarding the sizes of the two standard stimuli within the number of illuminations made by the normals.

3. Significant interactions between the standard stimulus conditions and the subjects indicated that the schizophrenics required significantly more illuminations in judging aide stimulus and made more errors than normal subjects in judging patient stimulus.

4. There was a positive correlation between number of errors and frequency of illuminations for both groups with aide as stimulus.

5. The absence of any differences in the size judgments of the two groups suggested that there is no deficit involved in size judgment of chronic undifferentiated schizophrenics. It is likely that size matching requires perceptual abilities which are primitive and fully developed prior to the initiation of a long term schizophrenic process. Difficulty in making judgments was accounted in terms of the schizophrenic's difficulty in attending to the relevant cues and rejecting the irrelevant cues within the stimulus field, consisting of both internal and external cue elements.

REFERENCES

ARIETI, S. *Interpretation of schizophrenia.* New York: Brunner, 1955.

BRUNER, J. S. Personality dynamics and perceiving. In R. Blake & G. Ramsey (Eds.), *Perception: An approach to personality.* New York: Ronald, 1951. Pp. 121–145.

COHEN, B. D. Motivation and performance in schizophrenia. *J. abnorm. soc. Psychol.,* 1956, **52,** 186–190.

COOPER, R. Objective measures of perception in schizophrenics and normals. *J. consult. Psychol.,* 1960, **24,** 209–214.

EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1958.

EISSLER, K. R. Notes upon defects of ego structure in schizophrenia. *Int. J. Psycho-Anal.,* 1954, **35,** 141–146.

FREUD, S. *The ego and the id.* London: Hogarth, 1947.

GOLDSTEIN, K. The significance of psychological research in schizophrenia. *J. nerv. ment. Dis.,* 1943, **97,** 261–279.

HERNANDEZ-PEON, R., SHERRER, H., & JOUVET, M. Modification of electrical activity in the cochlear nucleus during "attention" in unanesthetized cats. *Science,* 1956, **123,** 331–332.

HUSTON, P. E., & SHAKOW, D. Studies of motor function schizophrenia: III. Steadiness. *J. genet. Psychol.,* 1946, **34,** 119–126.

KLEIN, G. S. The personal world through perception. In R. Blake & G. Ramsey (Eds.), *Perception: An approach to personality.* New York: Ronald, 1951. Pp. 328–355.

LEIBOWITZ, H. W., & PISHKIN, V. Size constancy in chronic schizophrenia. *J. consult. Psychol.,* 1961, **25,** 196–199.

LOVINGER, E. Perceptual contact with reality in schizophrenia. *J. abnorm. soc. Psychol.,* 1956, **52,** 87–91.

LYDECKER, W. A., PISHKIN, V., & MARTIN, B. The effects of different feedback conditions on the concept identification of schizophrenics. *Psychol. Rep.,* 1961, **9,** 255–263.

MCNEMAR, Q. *Psychological statistics.* New York: Wiley, 1955.

PISHKIN, V. Effects of probability of misinformation and number of irrelevant dimensions upon concept identification. *J. exp. Psychol.,* 1960, **59,** 371–378.

PISHKIN, V. Transmission of information as a function of misinformation feedback distribution. *Psychol. Rep.,* 1961, **9,** 255–263.

PISHKIN, V., ARMSTRONG-RESSY, C. T., ALLER, R. A., & COMSTOCK, R. L. The effects of temperature on nerve excitability, attention, and reaction time in chronic schizophrenia. *J. nerv. ment. Dis.,* 1960, **131,** 348–353.

RABIN, A. I., & KING, G. F. Psychological studies. In L. Bellak (Ed.), *Schizophrenia.* New York: Logan, 1958.

RAUSH, H. L. Perceptual constancy in schizophrenia: I. Size constancy. *J. Pers.,* 1952, **21,** 176–187.

RAUSH, H. L. Object constancy in schizophrenia: The enhancement of symbolic objects and conceptual stability. *J. abnorm. soc. Psychol.,* 1956, **52,** 231–234.

SALZINGER, K. Shift in judgment of weights as a function of anchoring stimuli and instructions in early schizophrenics and normals. *J. abnorm. soc. Psychol.,* 1957, **55,** 43–49.

SANDERS, R., & PACHT, A. R. Perceptual size constancy of known clinical groups. *J. consult. Psychol.,* 1952, **16,** 440–444.

VOTH, A. C. An experimental study of mental patients through the autokinetic effect. *Amer. J. Psychiat.,* 1947, **103,** 793–805.

WECKOWICZ, T. E. Size constancy in schizophrenic patients. *J. ment. Sci.,* 1957, **103,** 475–486.

WECKOWICZ, T. E., & BLEWETT, J. Size constancy and abstract thinking in schizophrenic patients. *J. ment. Sci.,* 1959, **105,** 909–934.

WHITMAN, M. The performance of schizophrenics on social concepts. *J. abnorm. soc. Psychol.,* 1954, **49,** 266–271.

ZAHN, T. P. Acquired and symbolic affective value as determinants of size estimation in schizophrenic and normal subjects. *J. abnorm. soc. Psychol.,* 1959, **58,** 39–47.

(Received August 23, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 331–335

LACK OF CORRELATION BETWEEN THE CALIFORNIA PSYCHOLOGICAL INVENTORY AND HYPNOTIC SUSCEPTIBILITY[1]

ERNEST R HILGARD AND LILLIAN W. LAUER

Stanford University

Individual susceptibility to hypnosis can be shown to be reliably measured and quite stable; because such individual differences appear to be manifestations of personality it appears plausible that they should be correlated with at least some of the scores on a general multidimensional personality inventory. Our results, for a sample of 110 male and 106 female students, show essentially a lack of correlation between the scales (and the items) of the California Psychological Inventory (CPI) (Gough, 1957) and the Stanford Hypnotic Susceptibility Scale (SHSS) (Weitzenhoffer & Hilgard, 1959).

These results are not wholly unexpected. Weitzenhoffer (1958), using a sample of 200 student subjects at the University of Michigan, found only trivial correlations with the Guilford-Zimmerman Temperament Survey and the Cattell 16 PF Questionnaire; and Sector (1961), with 48 subjects, found no predictive significance within the Minnesota Multiphasic Personality Inventory (MMPI), either for established scales or by way of an item analysis. There are, however, occasional reports of more significant results with personality tests, both Sarbin (1950) and Faw and Wilcox (1958) reporting some positive correlations with the MMPI. More recently Furneaux and Gibson (1961) reported positive correlations between hypnotic susceptibility and the Maudsley Personality Inventory (MPI), but their sample was a curious mixture of 55 students and 44 older readers of a radical journal, with no breakdowns to tell what this heterogeneity produced. Lang and Lazovik [2] with 32 student subjects report some marginally significant positive correlations (.35, .36) with the social extraversion scale of the Pittsburgh revision of the MPI (Bendig, 1959), and with the affiliation scale of Edwards' PPS. The signs of the correlations with neuroticism and anxiety were negative, but the correlations were not significant. In general, when college student samples have been used, the results with the ordinary personality inventory have not been encouraging; our results simply add another sizable sample to the history of failure.

[1] A report from the Laboratory of Human Development, founded under a grant from the Ford Foundation. The research on hypnosis, begun under that grant, has been continued with the aid of the Robert C. Wheeler Foundation and the National Institute of Mental Health (Grant M-3859).

METHOD

During 2 academic years (1958–1959, 1959–1960) the CPI was routinely administered to all students enrolled in the introductory psychology course at Stanford; scores were obtained without any reference to the undertaking of experiments on hypnosis. Later an opportunity was given for participation in experiments on hypnosis as part of the outside laboratory requirement of the course, and a fairly representative sample of each class was obtained. The mean results for some of the CPI scales differed significantly from the published norms (Table 1) but the mean differences were not great in magnitude and could reflect differences between Stanford students and the general norm as well as differences related to signing up for hypnotic experiments.

The two forms of the SHSS were individually administered on two successive days under standardized conditions. The test is essentially a work-sample of hypnosis, involving a standard induction followed by a number of suggestions to which the subject may or may not respond. There are 12 scorable items on each form, so that the total 2-day score yields a score ranging from 0 to 24. The 2-day reliability, based on the test-retest reliability stepped up by the Spearman-Brown formula was .91 for the 124 subjects (62 male, 60 female) tested during 1958–1959 and .95 for the 96 subjects (48 male, 48 female) tested in 1959–1960.

Data analysis took two forms. One consisted of a correlation between the standard scales of the CPI

[2] Lang, P. J., & Lazovik, A. D. Personality and hypnotic susceptibility. Unpublished manuscript, 1961.

TABLE 1

HYPNOTIC SAMPLE COMPARED WITH CPI NORMS FOR COLLEGE STUDENTS

CPI scales	Norms for college men (N = 680)		Hypnotic sample (male) (N = 110)		Norms for College women (N = 2120)		Hypnotic sample (female) (N = 106)	
	M	σ	M	σ	M	σ	M	σ
Dominance	28.5	6.0	28.8	6.6	28.5	5.9	29.0	5.9
Capacity for Status	21.4	4.0	21.5	3.6	22.2	3.6	22.9*	3.0
Sociability	25.4	5.3	25.5	5.5	26.0	4.8	26.7	4.5
Social Presence	37.6	6.6	37.8	5.8	37.0	5.9	38.2*	5.9
Self-acceptance	20.6	6.3	22.8***	3.8	19.5	8.1	23.4***	3.7
Sense of Well-being	37.2	4.9	36.8	4.2	37.5	4.4	36.5**	4.3
Responsibility	31.5	4.7	31.3	4.6	33.3	4.1	32.3**	4.4
Socialization	37.5	5.2	37.0	6.1	39.5	5.0	38.7	5.2
Self-control	29.2	7.1	26.5***	7.3	30.8	7.4	28.3***	6.5
Tolerance	24.0	4.8	24.1	4.9	25.0	4.2	25.6	3.8
Good Impression	18.1	6.3	16.4***	5.8	19.1	6.2	16.7***	4.9
Communality	25.3	2.2	25.4	2.1	25.5	2.0	25.6	2.1
Achievement via Conformance	28.6	4.6	27.7	4.6	28.8	4.4	28.5	4.4
Achievement via Independence	21.6	4.1	22.7**	4.2	21.9	3.9	23.9***	4.0
Intellectual Efficiency	40.8	5.2	40.8	5.4	41.4	4.8	41.9	6.1
Psychological-mindedness	11.9	3.1	12.1	2.7	11.4	2.9	11.9	3.3
Flexibility	11.4	3.8	12.1	3.7	11.6	3.7	13.3***	3.8
Femininity	16.8	3.8	16.2	3.5	22.8	3.3	22.5	3.1

* Mean of hypnotic sample differs from that of normative sample at .05 level (t-test, two-tailed).
** Difference at .01 level.
*** Difference at .001 level.

TABLE 2

PRODUCT–MOMENT CORRELATIONS OF 2-DAY SHSS SCORES WITH CPI SCALES

CPI scales	Males N = 110	Females N = 106
Dominance	−.11	.23**
Capacity for Status	.07	−.02
Sociability	−.09	.12
Social Presence	−.02	.09
Self-acceptance	−.13	.25**
Sense of Well-being	−.20*	−.09
Responsibility	−.13	.08
Socialization	−.11	.09
Self-control	−.19*	−.12
Tolerance	−.11	.03
Good Impression	.00	−.02
Communality	−.18	.26**
Achievement via Conformance	−.07	.12
Achievement via Independence	.01	.04
Intellectual Efficiency	−.10	−.05
Psychological-mindedness	−.15	−.03
Flexibility	−.03	−.03
Femininity	.09	.15

* Significant at .05 level.
** Significant at .01 level.

and the SHSS, done separately for the men and for the women, and for the 2 years. The second was an attempt to build a new predictive scale based on an item analysis of the first year's scores, then validated on the scores of the following year. This also was done separately for men and for women.

RESULTS

Correlations with standard scales of the CPI. Pearsonian correlations between the 2-day SHSS scores and the CPI scale scores, computed separately for males and females, yielded the results in Table 2. While a few of the scales reach significance, none is significant for both men and women, and the significance is computed on the basis of the single correlations, i.e., does not take into account that when so many correlations are computed, a higher significance level should be demanded for any one that is to be interpreted as significant. Of the three scales significant for females over the two samples combined, none was significant for both samples separately and none was significant for the males on either sample. We therefore

interpret these results as essentially negative.

An attempt to derive a scale by item analysis. On the assumption that the item pool of 480 miscellaneous items in the CPI would almost surely have some experiences related to hypnosis, we proceeded on a purely empirical basis to select those items which, for the first year's sample, were differentially related to hypnotic susceptibility, a scale thus derived then to be validated against the next year's sample. By selecting items that discriminated between the upper half and lower half of hypnotic subjects, we were able to construct a 33-item scale for males that correlated $r = .67$ with their SHSS scores ($N = 62$), and a 36–item scale for females that correlated $r = .66$ with their SHSS scores ($N = 60$). These results, while promising, led to some initial skepticism because there were only three items common to both the male and female scales as thus empirically derived.

When the new scales were scored on the next year's sample, the correlations fell to a negligible level, for the males $r = -.05$ ($N = 48$), for the females, $r = +.08$ ($N = 46$). Thus the effort to construct a scale predictive of hypnosis on this purely empirical basis was a failure.[3]

DISCUSSION

Our results are consistent with the large body of previous results obtained in attempting to relate hypnotic susceptibility to scores on structured personality inventories. Occasional positive findings have been on small samples, unusual samples, or in studies that were not replicated.

Several possibilities can be suggested:

1. Perhaps hypnotic susceptibility is not a general personality characteristic, but is some highly specialized ability that can be studied only by sampling it. This we have sometimes whimsically called the "diabetes" theory: there it is, but it doesn't correlate with anything else. Some people merely inherit the tendency to diabetes; some are hypnotizable and some are not. We do not credit this theory very highly, for reasons to be given. It has to be acknowledged, however, and the fact that many distributions of hypnotic susceptibility turn out to be bimodal (Hilgard, Weitzenhoffer, Landers, & Moore, 1961), gives it a small measure of support. Even though the ability is specialized there are surely *some* nonhypnotic experiences that correlate with it.

2. Pehaps hypnotic susceptibility rests upon "deeper" or "more unconscious" aspects of the personality than are tapped by structured pencil-and-paper tests, and projective tests, such as the Rorschach or TAT would do better. The evidence (which we shall not review at this time) with respect to the Rorschach is not good, and with respect to the TAT is too fragmentary; whether or not some projective device might work, we are convinced that the distinction between pencil-and-paper and projective tests does not lie at the heart of the matter.

3. Despite the large number and variety of questions on inventories such as the CPI or the MMPI it is quite possible that they miss domains of experience relevant to hypnosis. Some work in our laboratory indicates that this is indeed the case. Building upon and extending some earlier work of Shor (1960) done elsewhere, Ås, O'Hara, and Munger (1962), working in our laboratory, drew up a pencil-and-paper test of 60 items emphasizing subjective experiences occurring outside of hypnosis that resemble those occurring in the hypnotic state: altered states of awareness, memory lapses, mystical experiences, change in the body image, role changes, etc. In the first test with a group of female subjects ($N = 50$), all but 12 of the 60 items correlated positively (though not always significantly) with hypnotic susceptibility.[4] The

[3] The fact that some existing scales correlated over the total sample means that an empirical scale could be constructed that would yield correlations at least that high within this total sample. Obviously by using the total sample, and capitalizing on chance, such a scale could be arranged, as indeed it was for our first sample. It would take additional replications to see what the residual correlations would be; present results suggest that they would be low.

[4] The measure of susceptibility he used, a scale under development in the laboratory known as the Consolidated Scale (CS), includes 7 items from the original Form B and 5 new items, tapping more cognitive functions, such as dreams and hallucinations. Scores on the CS correlate $r = .84$ with Form A of the SHSS ($N = 102$ female subjects), which is about as high as the correlation between the

overall score, based on a priori assigning of the positive and negative scores to the replies and scoring all items, yielded a significant correlation of $r = .36$ with the hypnotic score; item selection produced a 16-item scale correlating .68 with the sample from which the scale was constructed, and these same 16 items correlated $r = .29$ with the next sample ($N = 52$). Thus it is possible to find pencil-and-paper items that will yield moderate correlations with hypnotic susceptibility; it appears that the CPI and MMPI may be lacking in enough of the kinds of items required. Note that the correlations, while significant, are not large; it takes an appreciable number of items of appropriate kind to produce a prediction of hypnotic susceptibility.

4. A fourth possibility is that a more subtle kind of scoring is needed. On the assumption that more than one personality syndrome can lead to hypnotic susceptibility, it may be necessary to use one "sorting" variable and a second "dimensional" variable. This is, of course, the logic that Eysenck (1959) has been advocating, and its possible usefulness is illustrated in the study by Furneaux and Gibson (1961). Using a typical Maudsley procedure they first sorted their very heterogeneous sample of 99 students and radical-reading adults into 22 "liars" and 77 "honest" subjects. Because only 5 of the 22 liars fell in the hypnotic group, as compared to 44 hypnotizable among the total 99, a high lie score was used to classify a person as insusceptible. The advantage of a double-classification was well illustrated for the honest subjects. There was no correlation between extraversion and hypnotic susceptibility, and a low negative correlation ($r_{tet} = -.37$)[5] between neuroticism and hypnotizability for the 77 honest subjects. Yet when extraversion and neuroticism were combined, the correlation between neuroticism and hypnotic susceptibility rose to $r_{tet} = -.75$ for the 44 most extraverted subjects, although it became a nonsignificant positive correlation ($r_{tet} = +.15$)[5] for the remaining 33 introverted subjects. Thus a variable (extraversion), itself not correlating with hypnosis, was a useful sorting variable. We have attempted to combine some of the CPI scales in meaningful ways to predict hypnotic susceptibility but thus far without success; we have not, of course, exhausted the possibilities, which are numerous. Our appraisal of subjects by interview does make it appear plausible to us that some kind of sorting variable would be helpful when any other personality dimension is to be used in predicting hypnotic susceptibility. We shall report attempts along this line at another time.

original Forms A and B, so that his positive results cannot be attributed solely to the criterion used.

[5] These correlations, not in the original report, have been computed by us from the data given there.

SUMMARY

The California Psychological Inventory (CPI), administered before the opportunity to sign up for hypnotic experimentation, was correlated with scores on the Stanford Hypnotic Susceptibility Scale (SHSS), following administration of both Form A and Form B of the scale. The correlations, based on a 2-year sample of 110 males and 106 females, were so small as to indicate essentially a lack of correlation between the CPI scales and scores on the SHSS.

An empirical item analysis, resulting in CPI scales correlating with SHSS scores within the sample on which the analysis was based, failed to hold up in validation the next year.

Several hypotheses to account for the negative results were examined including (*a*) the possibility that hypnotic susceptibility is so specialized as not to lend itself to prediction by personality tests, (*b*) that pencil-and-paper tests are inappropriate to something as subtle as hypnotic susceptibility, (*c*) that the content of the items of the CPI does not sufficiently represent hypnosis-like experiences, and (*d*) that sorting variables may have to be used before significant correlations can be obtained. The first two of these conjectures are not defended, but the last two are.

REFERENCES

ÅS, A., O'HARA, J. W., & MUNGER, M. P. The measurement of subjective experiences presumably related to hypnotic susceptibility. *Scand. J. Psychol.*, 1962, **3**, No. 1.

BENDIG, A. W. College norms for the concurrent validity of the Pittsburgh revisions of the Maudsley

Personality Inventory. *J. Psychol. Stud.*, 1959, **11**, 12–17.

EYSENCK, H. J. *The structure of human personality*. London: Methuen, 1959.

FAW, V., & WILCOX, W. W. Personality characteristics of susceptible hypnotic subjects. *J. clin. exp. Hypnosis*, 1958, **6**, 83–94.

FURNEAUX, W. D., & GIBSON, H. B. The Maudsley Personality Inventory as a predictor of susceptibility to hypnosis. *Int. J. clin. exp. Hypnosis*, 1961, **9**, 167–177.

GOUGH, H. G. *California Psychological Inventory*. Palo Alto, Calif.: Consulting Psychologists, 1957.

HILGARD, E. R., WEITZENHOFFER, A. M., LANDES, J., & MOORE, R. K. The distribution of susceptibility to hypnosis in a student population: A study using the Stanford hypnotic susceptibility scale. *Psychol. Monogr.*, 1961, **75**(8, Whole No. 512).

SARBIN, T. R. Contributions to role-taking theory: I. Hypnotic behavior. *Psychol. Rev.*, 1950, **57**, 255–270.

SECTER, I. I. Personality factors of the MMPI and hypnotizability. *Amer. J. clin. Hypnosis*, 1961, **3**, 185–188.

SHOR, R. E. The frequency of naturally occurring "hypnotic-like" experiences in the normal college population. *Int. J. clin. exp. Hypnosis*, 1960, **8**, 151–163.

WEITZENHOFFER, A. M. Personality and hypnotic susceptibility. *Amer. J. clin. Hypnosis*, 1958, **1**, 79–82.

WEITZENHOFFER, A. M., & HILGARD, E. R. *Stanford hypnotic susceptibility scale*. Palo Alto, Calif.: Consulting Psychologists, 1959.

(Received September 25, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 336–339

INTELLECTUAL IMPAIRMENT IN SCHIZOPHRENIA

RICHARD M. GRIFFITH, BETSY WORTH ESTES,[1] AND SELWYN A. ZEROF[2]

Veterans Administration Hospital, Lexington, Kentucky

The routine data gathered on inductees into military service and the subsequent testing of some of them in Veterans Administration hospitals offer the rare opportunity to evaluate the impact of schizophrenia on intellectual functioning.

Making use of service scores, Mason (1956) and Miner and Anderson (1958) reach apparently contradictory conclusions as to a relationship between emotional adjustment in general and intellectual level; but both seem to agree that schizophrenia—at least some types of it—tends to strike those of less than average intelligence. The serviceman, presumably functioning adequately, who was later to become schizophrenic, scored lower than did the general inductee on the Army General Classification Test (AGCT) (The Adjutant General's Office, 1945). Aside from the question of the *preillness* intelligence, Griffith and Taylor (1960) obtained a correlation in line with correlations reported in normal populations between years of education and IQ *obtained after mental breakdown* in a group of hospitalized veterans including schizophrenics. They inferred from this tangential bit of evidence that the IQ scores of psychotic patients are more valid measures than clinicians have allowed. More direct evidence is needed on the question of whether patients caught in the throes of an apparently disruptive process do suffer an impairment of intellectual performance.

It was possible to bring together results on the AGCT with one of the Wechsler forms (I, II, WAIS) on 65 schizophrenic patients. Although, strictly speaking, the AGCT is not an intelligence test, it correlates highly with standard measures of IQ (in the neighborhood of .85, or as highly as does WAIS with WBI!) (Wechsler, 1958, pp. 105, 115) and thus, broadly speaking, is comparable. By reducing the scores to standard scores the comparison could be made between the relative position of each patient before illness with his position after illness in the respective standardization populations. In 53 of the cases the hospital test had been administered within the first month following admission while the service test had been regularly administered shortly after induction. Most of the patients had entered service after WWII; only 12 of them were inducted at the time when press of manpower shortage had forced a loosening of standards in literacy-intelligence and in emotional adjustment.

The average age of the patients at testing within the hospital was 29.8 (6.6) years (*SD*s given in parentheses). Educational level was 9.8 (2.9) years of schooling. The time elapsed between testings was lengthy as test-retest studies go, averaging 8.0 (4.6) years. Hospital IQ on Wechsler's scales was 92.2 (13.8); the deficiency from the national norm may (*a*) reflect the norm of the geo-cultural area from which the patients came, or may (*b*) support the position that the less intelligent are prone to schizophrenia, or may (*c*) signify that schizophrenia impairs intellectual functioning. The latter possibility is the one which the present study investigates.

RESULTS

Several correlations leading to the crucial comparison may be presented. The correlation between the hospital IQ and years of education was lower, though not quite significantly so, than the previous Griffith and Taylor figure on a more heterogeneous sample of patients in the same hospital, .48 vs. .62. Service scores and education level correlated

[1] University of Kentucky, Consultant to Veterans Administration Hospital.

[2] Credit for this study should be shared with the personnel of the United States Army Records Center; St. Louis, Missouri, Eugene S. Tarr, commanding. Within the hospital we were assisted in the tedious combing of files and the collating of information by several voluntary workers: Jane S. Wiedeman, Vivian H. Taylor, Reubena Bailey, and George J. Haus. The only reward to any of these many donors is the inadequate recognition in this small footnote.

.64 (in line with values obtained on the run of inductees) (The Adjutant General's Office, 1945, p. 765). Hospital IQs and service scores correlated .70 to be compared with the .85 mentioned above as reported by Wechsler for normal persons and which foretells the answer to the central question.

A distribution of differences—gains and losses—between individual service and hospital standard scores was drawn up. The standard scores (z's) were obtained using a *SD* of 15 for IQ scores and a *SD* of 20 for the AGCT scores which are reported in IQ-like figures (around a mean of 100). The mean of this distribution of changes in scores was a plus .195 (.75) representing a net *gain in intellectual performance under the condition of schizophrenia*. This gain, equivalent to about three Wechsler IQ points, is significantly different at the .05 level from the neutral .00.

However, within this distribution, which on the average shows a gain in intellectual functioning, may be found the single instance to prove that the schizophrenic may give an invalid test. The over-all increase is pulled down by one patient who showed an extreme loss, a difference of 2.85 in the two z scores or more than twice as much as the next largest loss. With the *SD* of .75 in the distribution of difference scores it may be calculated that such a drop would have been exteremely improbable by chance. His record reveals that he was severely suspicious and inhibited at time of testing. He repeated three digits forward and none backward. He perceived nothing missing on any of the Picture Completion cards. In addition, he earned a zero score on Object Assembly. He rambled about "heaven" during the Picture Arrangement subtest. Quite correctly, the clinician noted that ". . . there are indications that his usual (intellectual) level is considerably higher. Illness has made definite inroads in this area." With this veteran who had completed 2 years of college the test doubtlessly had value as a projective instrument.[3]

[3] Our original sample of service scores included 32 on the AFQT (Armed Forces Qualification Test) (Uhlaner, 1952) which replaced the AGCT at the time of the unification of the Services and which was designed to screen prospective inductees in order to eliminate the unfit and to assign those who qualified fairly among the Services. However, as was pointed out to us, there was a period from 10 July 1950 to 1 December 1951 when careful standardization came to nought. The test was released to induction and recruiting offices where it was administered under conditions which resulted in what came to be termed "operational slippage"; particularly, the scores fell into an almost comically bimodal distribution with a pylon peak just on the high side of the point of disqualification. New "conversion" tables were drawn up but they proved ineffective and the unfortunate situation was not corrected until steps were taken to insure standard conditions in a relatively few testing centers. Twenty-three of our sample of 32 had taken the AFQT under the nonstandard conditions, and, interestingly enough, 11 of these earned scores were very near to the cut-off point. We discarded the whole body of data—though it had been hard to come by.

Taking the questionable scores on face value the group showed some loss of functioning in the hospital ($-.12$ average difference in z scores which became $-.04$ with one grossly discrepant score omitted, neither value approaching significance). Those 9 with legitimate AFQT scores (who took the test after 1951) showed a net loss of .26 in the hospital which is approximately the same value which the complete sample would yield if the illegitimate scores were "reconverted" as though they had been collected under standard conditions. If this discrepancy is representative it would signify a decrement equivalent to about 4 Wechsler IQ points in the schizophrenic.

DISCUSSION

However challenging to speculate far afield about the source of the unexpected result, one explanation is close at hand and prosaic. It may be found in a peculiarity in the original standardization of the AGCT. For some uncertain reason the more than 8,000,000 veterans who were administered the AGCT in service (there were eventually several forms of it) scored higher than did the 2,675 upon whom it was standardized (The Adjutant General's Office, 1945, p. 763). It may be gathered that the field mean was roughly four AGCT points higher which may be converted to an estimated z of .20. This correction would wipe out the gain found in the present study.

The very foundation for the study is an assumed equivalence between the standardization populations of the Wechsler and the service tests. Any conclusions from an examination of the published descriptions of the compositions of the populations would be farfetched.

The two mean points float independently of one another until the scales are empirically tied together; one test may be inherently "easier" than the other. Tamminen (1951) summarizes the results of back-to-back administration of the AGCT and Wechsler on a sample of 100 nondisabled veterans selected from six vocational counseling centers in Minnesota. Though the IQ level was above average (Wechsler IQ, 109.5) the group would seem to be a reasonably good control for the subjects in the present study. Their mean Wechsler reduces to a z of +.63, the mean AGCT to a z of +1.01. Thus, if anything, the AGCT was the easier of the two tests; the difference, equivalent to 5.7 IQ points, was in the direction to accentuate the present findings.

Many factors would make the results more conclusive had it been possible to control them. For example, it was not practical to track down previous hospitalizations in military, state, or other VA hospitals and impossible to determine whether the patients might have received a Wechsler earlier. That the gain in functioning in the hospital was not due to any possible practice effect, however, may be concluded from a comparison of the relative Verbal and Performance scores of the patients who gained in the hospital vs. those who lost. The Performance sections of the Wechslers are much more subject to practice than the Verbal sections (Wechsler, 1958, p. 157). Yet, those who contributed to the gain in the present study showed a wider spread between Verbal and Performance (in favor of the Verbal) than did those who did not. As another example, a small but statistically significant Pearson r—minus .26—between education and gain in functioning raises questions which cannot be definitely answered. Eliminating the previously mentioned schizophrenic with 2 years of college who gave a Wechsler IQ of 61 drops this correlation somewhat below significance, but the fact may be that the AGCT penalized the poorly educated who later became patients. Again, who can estimate the effect on test scores of the enriching experiences in service for those with deprived backgrounds? (The four veterans who had completed the fifth grade or less averaged a gain of .95 in z scores, i.e., about 14 IQ points.)

But perhaps it is too easy to degrade a valid finding because it was unexpected. The possibility of a gain in IQ under acute schizophrenia has been broached in the literature: Fitzherbert (1955) describes dramatic increases in three young girls. It is conceivable that some schizophrenics should function more effectively after breakdown; for them, schizophrenia would represent a resolution of tensions, inhibitions (or whatever) which would impede performance. Particularly might induction into service and the circumstances usually surrounding service testing (in groups, etc.) disturb the preschizophrenic. Though the data are too incomplete and inconclusive to report at this time, it is suggestive that some of those schizophrenics in the present study who were doomed to remain in the hospital for the longer periods tended to gain in hospital vs. service scores more than did those who recovered most quickly and permanently. Thus, factors adversely affecting test performance may have favorable prognostic implications. The whole matter can only be left suspended until further research reduces the scientific tension.

In conclusion, however dubious the evidence for a gain in intellectual performance contingent upon a schizophrenic break, at least there is little evidence in these data for any material loss in performance, certainly not one which would be clinically applicable to the individual.

By all logic, tests which require concept formation, language, perceptual accuracy, attention, concentration, a clear sensorium and, particularly, motivation, must suffer in radical schizophrenia. Beyond any reasonable question the mute, the rigid, the incoherent would perform poorly (as perhaps the case cited above may illustrate). Thus, the above results can only be generalized to those newly admitted schizophrenics whom clinicians generally regard as testable.

The present study supports the evidence of the earlier (Griffith & Taylor, 1960) and reiterates that conclusion: "A large number of diagnostic protocols surveyed in the present study asserted on the basis of one bit of evidence or another that the patient's intelli-

gence was probably higher than the measured value. It would seem that very often this qualification was ill-founded."

SUMMARY

The intelligence quotients of veterans hospitalized for schizophrenia were compared with their scores on the Army General Classification Test—which they had taken some years before when, presumably, they had been functioning adequately. Contrary to all expectations, there was a net *gain* in level of intellectual performance after the onset of schizophrenia. Looking ahead in the possibility that these results which would explode some current concepts may be substantiated by further work, a bold rationale is timidly extended.

REFERENCES

FITZHERBERT, JOAN. Increase in intelligence quotient at onset of schizophrenia: Three adolescent cases. *Brit. J. med. Psychol.*, 1955, **28**, 191–193.

GRIFFITH, R. M., & TAYLOR, VIVIAN H. Effect of mental illness on intelligence test scores. *J. clin. Psychol.*, 1960, **16**, 352.

MASON, C. F. Pre-illness intelligence of mental hospital patients. *J. consult. Psychol.*, 1956, **20**, 297–300.

MINER, J. B., & ANDERSON, J. K. Intelligence and emotional disturbance: Evidence from Army and Veterans Administration records. *J. abnorm. soc. Psychol.*, 1958, **56**, 75–81.

TAMMINEN, A. W. A comparison of the Army General Classification Test and the Wechsler Bellevue Intelligence Scales. *Educ. psychol. Measmt.*, 1951, **11**, 646–655.

THE ADJUTANT GENERAL'S OFFICE, Classification and Replacement Branch, Personnel Research Section Staff. The Army General Classification Test. *Psychol. Bull.*, 1945, **42**, 760–768.

UHLANER, J. E. Development of Armed Forces Qualification Test and predecessor Army Screening Tests, 1946–1950. *Personnel Res. Branch Rep.*, 1952, No. 976.

WECHSLER, D. *The measurement and appraisal of adult intelligence*. Baltimore: Williams & Wilkins, 1958.

(Received October 5, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 340–344

THE EFFECT OF ORDER OF PRESENTATION ON PERCEPTION OF TAT PICTURES[1]

ADELAIDE DOLLIN

Institute of Living, Hartford, Connecticut

AND JAMES M. SAKODA

University of Connecticut

The investigation described here was concerned with the application of Helson's (1955) adaptation-level theory to projective test stimuli. Murstein (1959), pointing out the susceptibility of projective test responses to variables other than personality, has suggested that such variables be investigated in the context of what Helson calls the Stimulus × Background × Organism interaction. He refers here to the theory which views response as resulting not only from factors internal to the organism but from the complex interactions among the subject, the stimulus with which he is confronted, and the total context in which the psychological examination or experimental task occurs. Using TAT pictures which varied in emotional tone or mood, the present investigation sought, in accord with Helson's theory, to observe the fluctuations of perceived mood as pictures were presented in different contexts.

The study was divided into two parts. In Part I the stimulus value of the TAT cards with respect to mood was assessed for the college population to be used. In Part II the cards were administered in different orders. Hypotheses were as follows:

1. Significant differences exist among TAT cards as well as among individuals. A difference among cards reflects the fact that judgments about mood are not purely subjective but that the pictures themselves have an objectively discernible mood. The hypothesis is predictable from the work of Eron (1950) and Murstein (1958).

2. Judgments about cards vary as their position in a series is changed. The implication here is that stimulus value and individual propensities alone do not determine response to the card but that context is also a factor.

3. Direction of changes in judgment can be predicted by assuming that, according to Helson's adaptation-level theory, a contrast effect operates.

Such a result is anticipated from the theory that response is made relative to adaptation-level. All stimuli affect adaptation level which varies with varying stimulation. The presentation of sad-looking pictures depresses the adaptation level and they subsequently appear less sad. In comparison to the depressed adaptation level or neutral point, happy pictures are then rated happier. The converse would be true of happy scenes.

METHOD

Part I

Fifty-one subjects, 25 men and 26 women, undergraduate volunteers from the University of Connecticut, participated in the first part of the investigation. All 30 TAT cards, excluding Number 16, the blank one, were administered to the subjects individually and rated by them for mood on a 5-point scale ranging from −2 to +2. Scale points were labeled sad, slightly sad, neutral, slightly happy, happy.

Between subjects the cards were thoroughly shuffled so as to eliminate the influence of systematic order effects on the ratings; thus each subject saw the 30 cards in a different order. The subject was presented with a pencil, data sheet, and rating scale, and with the stack of TAT cards face down. The following instructions were given him.

> I'm interested in people's opinions about these pictures. Please look at each one and rate it on this scale. If you think the card on the whole has

[1] The paper is based on a doctoral dissertation completed by the senior author under the guidance of the junior author and submitted to the University of Connecticut in partial fulfillment of requirements for the PhD degree. The senior author would like to acknowledge the assistance of Sam L. Witryol and Jesse G. Harris, Jr., doctoral committee members, and of Marvin Reznikoff, Institute of Living.

The authors are also indebted to the Massachusetts Institute of Technology Computation Center for assistance in data processing.

a happy mood, rate it +2, if you think it has a sad mood, rate it −2 . . . (and so on for each of the scale's points). Look at the cards in the order you find them here—top one first, bottom one last. Write the number of the card on this sheet, this number here (pointing), and put your rating down next to it. Any questions?

Part II

Of the 30 TAT cards, 21 were selected; 7 rated happiest (6GF, 8GF, 9BM, 10, 12BG, 13G, 17BM); 7 saddest (3GF, 3BM, 6BM, 12M, 15, 18GF, 18BM); and 7 intermediate (1, 5, 7GF, 7BM, 11, 13B, 19). The 21 cards were chosen with two criteria in mind, (*a*) to maximize the difference between the three mood categories, (*b*) to minimize the difference between sexes. The seven cards of a given mood were considered one "treatment" and are referred to as such hereafter.

One hundred twenty subjects, 60 males and 60 females, also undergraduate volunteers, participated in Part II of the investigation. Each rated all 21 cards which were administered in one of six different orders (H, happy; N, neutral; S, sad): HNS, HSN, NSH, NHS, SNH, SHN. There were 20 subjects, 10 men and 10 women, in each of these six groups.

Within the limits of this design subjects were assigned to the groups at random. The seven cards comprising each treatment—H, S, or N—were given in 13 different random orders so as to eliminate systematic within treatment order effects.

In view of the results of Part I, the rating scale was changed to an 8-point one with the points labeled very, very sad, sad, slightly sad, neutral, slightly happy, happy, very happy. Instructions to the subjects were as follows:

> I'm interested in people's opinions about these cards. Please look at each one as I hand it to you and rate it on this scale. If you think the card on the whole has a very happy mood, rate it 8 . . . (etc.). Number each card, first one one, second one two, and so forth, and put your rating down next to that number. Any questions?

The subject was handed each card in order to control the length of time he looked at it (at least 5 seconds), and to make sure that the predetermined sequence was followed. Any suggestion that there were three different kinds of cards or a predetermined order was avoided.

TABLE 1

ANALYSES OF VARIANCE OF TAT CARD MOOD RATINGS IN A COLLEGE POPULATION

Source	*df*	*MS*	*F*
	Males		
Pictures	29	20.17	22.37***
Subjects	24	1.71	1.89**
Discrepance	696	.90	
	Females		
Pictures	29	22.36	51.29***
Subjects	25	2.48	5.69***
Discrepance	725	.44	

** Significant at .01 level.
*** Significant beyond .001 level.

TABLE 2

RATINGS OF TAT CARDS OF VARYING MOOD BY DIFFERENT ORDER GROUPS

Group	Happy	Neutral	Sad	Group mean
HNS	5.26	3.96	2.49	3.91
HSN	5.45	4.64	2.76	4.28
NHS	5.86	4.48	2.49	4.28
NSH	5.95	4.21	2.78	4.31
SHN	5.76	4.60	2.94	4.43
SNH	5.80	4.71	2.92	4.48
Treatment mean	5.68	4.43	2.73	4.28

Note.—Figures are means for 20 subjects and 7 pictures with a possible range of from 1 (very, very sad) to 8 (very happy).

RESULTS

Part I

Analyses of variance, presented in Table 1, indicated that there were significant differences among TAT cards as well as among subjects. From these results it can be inferred that both the stimulus value of the card and the predispositions of the subject contribute to produce the rating assigned.

Subject's ratings also indicated that the TAT pictures were, in general, gloomy in emotional tone. There was no systematic difference between men and women on the cards designed especially for men or women. Although, as previously reported in the literature (Eron, Terry, & Callahan, 1950; Newbigging, 1955), females seemed to assign sadder ratings, with a mean of −.50 as compared to males −.35, a *t* test indicated that the difference was not significant.

To compare picture rating with story rating as a method, data obtained in Part I by picture rating were compared to those of Eron, Terry, and Callahan (1950), who had complete TAT stories rated for mood. Ranks were assigned to the 19 cards used in common and rho was computed. The values were .81 for men and .76 for women, both of which were significant beyond the .001 level.

TABLE 3

SPLIT PLOT ANALYSIS OF VARIANCE OF ORDER GROUPS AND TAT CARD TREATMENTS

Source	*df*	*MS*	*F*
Main plot analysis			
Order groups	5	16.93	3.96**
Subjects within Order groups	114	4.27	
Subplot analysis			
Treatment	2	1845.26	778.59**
Treatment × Order groups	10	5.08	2.14*
Subplot error	228	2.37	

* Significant at the .05 level.
** Significant beyond the .01 level.

Part II

The subjects' ratings of H, S, and N pictures for the six order groups are given in Table 2. It should be noted that the labels H, S, and N are not accurate in an absolute sense since, for example, N cards are slightly sad. However the labels were retained as a convenient method of referring to the happiest, saddest, and intermediate cards of the TAT series.

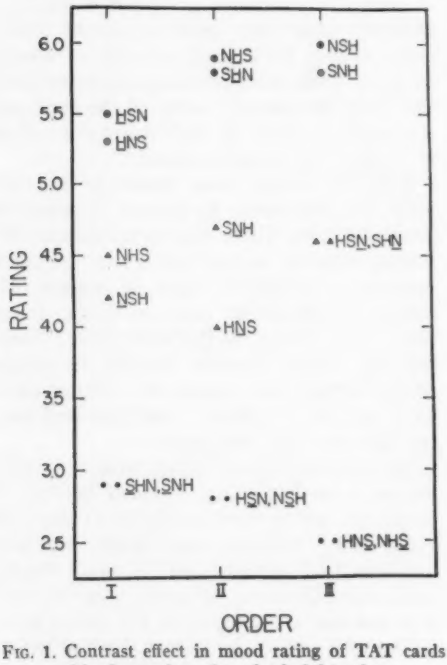

FIG. 1. Contrast effect in mood rating of TAT cards with change in order of administration.

An analysis of variance, presented in Table 3, indicated that there was indeed a significant difference among the three treatments, H, S, and N, a finding which confirmed the results of Part I concerning the effect of stimulus differences on response. Differences among the means of the six order groups were also found to be statistically significant. Since all groups received the same treatments, the observed difference among them, as well as the significant Treatment × Group interaction, can be attributed to order effects.

Inspection of the various means for the treatments in different order positions reveals trends in the hypothesized directions for S and H (Figure 1). These trends are statistically significant (see Table 4) and can be described in terms of contrast effect. When H, for example, is given second it has been preceded by sadder cards; hence it seems happier than it did when seen first.

TABLE 4

ANALYSES OF TRENDS IN MOOD RATING OF TAT CARDS WITH CHANGE IN ORDER OF ADMINISTRATION

Source	*df*	*MS*	*F*
Treatment H			
Between groups	5	9.88	
Linear regression	1	38.55	19.71*
Error of estimate	4	2.99	1.66
Treatment N			
Between groups	5	11.67	
Linear regression	1	10.58	4.56*
Error of estimate	4	11.94	5.15*
Treatment S			
Between groups	5	5.54	
Linear regression	1	27.02	5.51*
Error of estimate	4	.17	0

* Significant at, at least, .05 level.

The case of the N treatment is more complex. Analysis revealed a curvilinear trend which was statistically significant. There was, in addition, an overall upward trend, also significant. The curvilinearity is the result of the averaging of the two N treatments given second. Inspection of the means separately shows, in accordance with the principle of contrast, that N cards seen after H are rated sadder than those seen first. Conversely, N

cards seen after S are rated happier. It appears also that neutral cards are rated happier when preceded by both S and H than they are at the first position. Consideration of the mean values of S and H cards combined, −.85, reveals that together they represent a sadder experience than do the N cards themselves whose mean is −.35. Thus, in contrast, N cards following SH seem happier.

DISCUSSION

Considering the results in the light of Helson's hypothesis concerning the determinants of response, Part I reflects stimulus influences which, although recognized, may still be undervalued in the case of projective tests. Organismic or personality factors have traditionally been seen as central and crucial in affecting response. As previous studies have reported, the TAT series is sad in mood rather than neutral. The finding that women tell sadder stories it not confirmed here in the case of ratings. However, tests of significance of the difference were not previously reported. Despite the lack of significance in the present investigation, the consistency with which a small discrepancy between males and females is found suggests that it is perhaps real although not statistically significant in a given instance.

In comparing this investigation with those of Eron and others who were also concerned with emotional tone, it should be borne in mind that in the past data have been collected by obtaining complete TAT stories from the subjects and then employing judges to rate the stories along the desired dimensions. In this study direct rating of the TAT cards themselves was the method used. A question then arises concerning the comparability of the two methods; the high correlation obtained between Eron's data and those of this study suggest that for the mood variable at least, the different methods yield similar results.

Stimulus and organismic factors having been demonstrated with respect to the mood variable, the main point of the study, interaction of stimulus and background influences, remains to be considered. Helson, who initially explored adaptation-level effect in the context of psychophysical experiments, suggested that the concept could appropriately be applied to more complex behavior. It does appear to be a significant factor when stimuli are the complex pictorial ones of the TAT series. The perceived mood of a picture depends not only upon its stimulus quality and the personality characteristics of the subject. The total experience represented by an array of cards ought also to be considered. The subject's evaluation of the stimulus, in other words, is not an isolated event; it is influenced by preceding experience with related stimuli. The results are in accordance with Helson's adaptation-level theory.

It is suggested that in the clinical situation when cards are administered in standard order an adaptation-level effect will occur. Since the cards are, on the whole, gloomy, those seen toward the end of the series will be judged happier than those seen first. To correct this trend a better balance between happy and sad cards would have to be achieved. It is proposed that the effect could be minimized by administering cards in alternate happy-sad order.

SUMMARY

The investigation described here was concerned with the application of Helson's adaptation-level theory to projective test stimuli. It has been suggested that such a theory would provide a framework within which the roles of stimulus, background and organismic factors in projective test response might be evaluated and organized.

It was concluded that, for the Thematic Apperception Test, both the stimulus quality of the pictures and the context in which they are seen, as well as organismic factors, are significant in influencing response. Pictures received significantly different ratings on the dimension of mood or emotional tone. In addition, perceived mood of the pictures varied significantly when they were presented in different contexts. In accord with Helson's theory, the direction of changes in perception can be anticipated by assuming that an adaptation-level effect will operate. For example, sad looking pictures were rated sadder when preceded by happy pictures.

It was suggested that the effect would op-

erate in the clinical situation when pictures are administered in standard order. As a consequence of the relative sadness of the pictures, one might expect to find adaptation to sadness. Thus, cards seen toward the end of the series would evoke happier stories than those seen earlier. The theory suggests that the effect could be minimized by noting the stimulus values of the pictures and administering them in alternate happy-sad order.

REFERENCES

ERON, L. D. A normative study of the Thematic Apperception Test. *Psychol. Monogr.*, 1950, **64**(9, Whole No. 315).

ERON, L. D., TERRY, DOROTHY, & CALLAHAN, R. The use of rating scales for emotional tone of TAT stories. *J. consult. Psychol.*, 1950, **14**, 473–478.

HELSON, H. An experimental approach to personality. *Psychiat. res. Rep.*, 1955, **2**, 89–99.

MURSTEIN, B. I. Nonprojective determinants of perception on the TAT. *J. consult. Psychol.*, 1958, **22**, 195–198.

MURSTEIN, B. I. A conceptual model of projective techniques applied to stimulus variations with thematic techniques. *J. consult. Psychol.*, 1959, **23**, 3–14.

NEWBIGGING, P. LYNN. Influence of a stimulus variable on stories told to certain TAT pictures. *Canad. J. Psychol.*, 1955, **9**, 195–206.

(Received October 11, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 345–350

RELATIONSHIPS AMONG THREE MEASURES OF CONFLICT OVER HOSTILITY[1]

J. THOR NELSON AND SEYMOUR EPSTEIN

University of Massachusetts

In a recent paper (Epstein & Fenz, 1962), a model for the measurement of approach-avoidance conflict with specially constructed projective techniques was presented. Essentially the model represents a modification and further development of Miller's (1951) theory of displacement. Following Miller, it is assumed that under certain conditions the gradient of avoidance is steeper than the gradient of approach. However, unlike Miller, the gradients are represented as hypothetical constructs of drive, defined as a directed force, rather than as responses. By algebraically summing the positive and negative gradients, a response measure of net approach increment is obtained, which corresponds to Miller's response measure of net approach. Two other indices of conflict are obtained for which Miller can not be held responsible, both of which follow from the conceptualization of the gradients as drives. First, the gradients are summed without regard to algebraic sign in order to derive gradients of activation, which it is assumed can be measured by physiological correlates. Secondly, it is assumed that performance is related to activation by an inverted U shaped function (Malmo, 1959), so that conflict, when of sufficient magnitude, can be expected to produce a decrement in level of performance. In summary, from the one basic model, the following three signs of conflict are derived for use with stimulus dimensions built into projective techniques:

1. An increment in net approach strength to stimuli of low relevance and a decrement in net approach strength to stimuli of high relevance, as measured by content of verbal response.

2. A positive gradient of physiological activation as a function of increasing stimulus relevance, as measured by the galvanic skin response (GSR).

3. A positive gradient of cognitive deficit as a function of increasing stimulus-relevance, as measured by goodness of response, or by reaction time in a test where speed is a factor.

It follows that subjects demonstrating conflict on any one of the above indices should differ on the others. The purpose of the present investigation is to ascertain whether the three indices of conflict are of sufficient reliability, validity, and generality, so that subjects selected on one of the indices on a particular projective technique demonstrate corresponding responses on the other indices on a different projective technique at a different time. More specifically, the study investigates gradients of GSR and reaction time on a word association test as related to verbal approach and avoidance responses on a TAT-type test, with a 6-month interval between sessions.

METHOD

Subjects

All subjects were male undergraduates at the University of Massachusetts. They were chosen by their responses to a TAT-like test in a previous experiment (Saltz & Epstein, in press) in which 180 student volunteers participated. The TAT-like test consisted of six specially designed pictures, two of which presented no hostile cues, but for which a hostile response was appropriate (low stimulus relevance), and one of which depicted a hostile scene (high stimulus relevance). The three critical pictures were:

1. Low stimulus relevance—A young man with an ambiguous expression is looking toward the doorway of a house.

2. Low stimulus relevance—A dimly illumined figure of a man is striding in a determined fashion from one room to another.

3. High stimulus relevance—A young man is holding an older man by the throat and pushing him against the railing of a staircase. The older man's eyes are closed, and he is slumped over. This picture

[1] This study is part of a project on the measurement of drive and conflict administered by Epstein and supported by Grant M-1293, National Institute of Mental Health, United States Public Health Service.

is identical to TAT picture 18 GF (Murray, 1943) with males substituted for females.

Stories were scored for the presence or absence of hostility on the part of the hero. On the basis of this criterion, two groups were selected. Group I, the "Conflict Group," consisted of those individuals who gave a hostile response to a picture of low hostility, and failed to give such a response to the picture of high hostility. Group II, the "Control Group," behaved in a manner that was more determined by the properties of the stimuli in that they did not produce a hostile response to the pictures of low hostility and did to the picture of high hostility. At a later date in the planning of this study it was decided to introduce two additional groups to determine whether any differences found between the first two groups were a function of the high or low relevant pictures alone. Accordingly, Groups III and IV were added: Group III, the "Uniform Expression Group," gave hostile responses to both the low and high hostility relevant pictures; Group IV, the "Uniform Inhibition Group," did not give hostile responses to either. Thus, there were four groups constituting a 2×2 design for presence and absence of hostility in regard to the response and the stimulus. Each of the groups was composed of eight subjects.

There was a 6-month time interval between initial testing and later testing. The preselected subjects were personally contacted and paid for their participation. Not one refused, thereby eliminating selective bias.

Materials

A 41-word association list was constructed which contained four words at each of three levels of relevance to hostility. The words were selected from a larger pool which had previously been sorted into three levels of increasing hostility by five clinical psychologists. Agreement was unanimous on all words selected. All were verbs referring to physical activities. The final list consisted of eight practice words followed by twelve randomly distributed experimental words, each of which was buffered by two neutral words. Following is the list of words, with the lowest level, or neutral words, indicated by an (N), the medium level by an (M), and the high level by an (H): dog, salt, quiet, red, joy, slow, citizen, music, *smash* (H), sky, table, *jump* (N), bitter, carpet, *push* (M), book, moon, *kick* (H), loud, taxi, *swim* (N), bone, swift, *chase* (M), radiator, green, *skate* (N), flower, wagon, *kill* (H), radio, woman, *climb* (N), smooth, yellow, *grab* (M), comfort, butter, *stab* (H), wind, sand, *tease* (M).

The words were presented by a tape recorder at 30-second intervals. The list was first screened for clarity of pronunciation by three judges, and adjustments were made until all words were correctly reported by all judges. A ready signal preceded each word by 2 seconds. Verbal responses were written by the experimenter, and reaction time from the end of the stimulus word to the beginning of the response was recorded by the experimenter who controlled a pen marker on the GSR record.

A Hunter Model 100A GSR amplifier was used for the measurement of GSR. The apparatus was adapted for use with finger electrodes, which were placed on the ventral tip of the index and middle fingers of the dominant hand. Prior to taping on the electrodes, the area of contact was washed with alcohol, and a layer of paste composed of Bentonite, glycerine, and Ringer's solution was applied to reduce polarization. In order to measure resistances of greater than 50,000 ohms, parallel resistors were placed in the circuit. When this was done conversion charts were used to obtain corrected readings.

Procedure

Upon entering the experimental room, the subject was seated facing the experimenter, at a small table. A large screen hid the equipment from view. About 3 minutes after the electrodes were applied, basal level was taken, and the following instructions given:

> Now we will begin a test of speed of reaction to words. After you have heard a word—and be sure that you have heard the whole word—say the first word that occurs to you as quickly as possible. If you are not sure you heard a word correctly, respond to what you thought it was. Please do not make any comments or ask any questions between the words, but save them for the end of the test. Now, any questions before we begin?

RESULTS

Reaction Time

In Figure 1 it can be seen that for the Conflict Group there is an increase in reaction time from neutral to medium hostility words, followed by a decrease to high hostility words. The inverted V shaped curve is further emphasized when the reactions of the Conflict Group are interpreted relative to those of the Control Group. Analysis of variance (see Table 1) reveals that the groups differ significantly (.05 level) in their reaction to the stimulus dimension. Individual analysis of each group indicates that while the Conflict Group responds significantly (.05 level) differently to the three levels of stimulus relevance, the Control Group does not.

An analysis of variance of mean reaction time of the four groups arranged in a 2×2 design for performance on pictures of low and high relevance to hostility indicates that only performance on the picture of high relevance is significantly related to differential reaction time along the stimulus dimension (.02 level). Inspection of the group means reveals that the Uniform Inhibition Group

produces a similar but more exaggerated inverted V shaped curve than the Conflict Group. Mean reaction times in order of increasing stimulus relevance are: Conflict Group: 1.77 seconds, 2.17 seconds, 2.04 seconds; Uniform Inhibition Group: 1.67 seconds, 2.17 seconds, 1.78 seconds. The Uniform Expression Group behaves in a similar manner to the Control Group, with mean reaction times in order of increasing stimulus relevance, as follows: Uniform Expression Group: 1.72 seconds, 1.49 seconds, 1.74 seconds; Control Group: 1.61 seconds, 1.57 seconds, 1.96 seconds. Inspection of the individual curves of reaction time confirms the analysis of the means.

TABLE 1

ANALYSIS OF VARIANCE OF REACTION TIME

Source of variance	*df*	*MS*	*F*
Between subjects	15		
Groups	1	1.33	3.91
Subjects	14	.34	
Within subjects	32		
Stimulus Relevance (SR)	2	.09	1.12
Groups × SR	2	.36	4.50*
Subjects × SR	28	.08	

* Significant at .05 level.

Galvanic Skin Response

As with reaction time, scores consist of the average response to the four words at a particular level in the stimulus dimension. A response is recorded as the increase in conductance, measured in micromhos, from the level immediately preceding the stimulus word to the highest point following it and not preceded by a change in direction.

Inspection of Figure 2 reveals that the Conflict Group demonstrates a marked increase in GSR from the neutral to the medium words followed by a decrease to the high

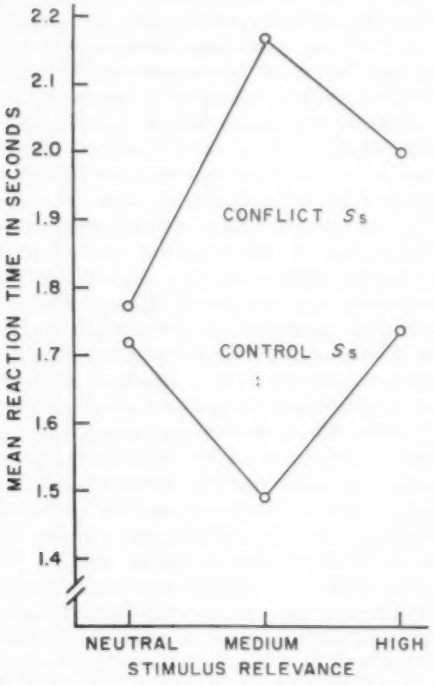

FIG. 1. Reaction time for Conflict and Control groups as a function of stimulus relevance.

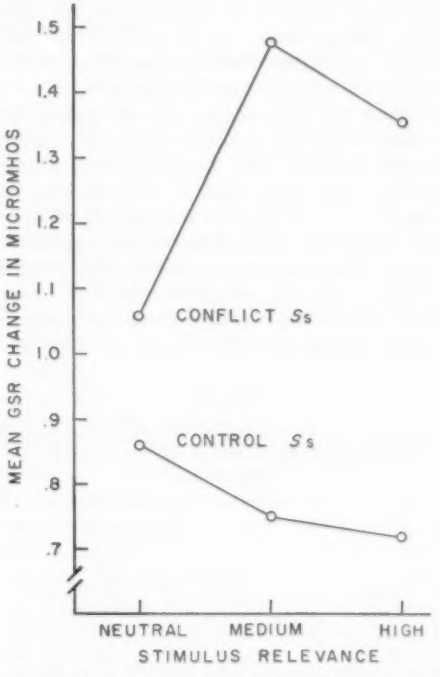

FIG. 2. GSR for Conflict and Control groups as a function of stimulus relevance.

TABLE 2

ANALYSIS OF VARIANCE OF GSR

Source of variance	*df*	*MS*	*F*
Between subjects	15		
Groups	1	323.18	3.61
Subjects	14	89.50	
Within subjects	32		
Stimulus Relevance (SR)	2	9.82	1.10
Group × SR	2	33.13	3.42*
Subjects × SR	28	9.66	

* Significant at .05 level.

hostile words. The Control Group, on the other hand, shows little change in GSR as a function of the stimulus dimension. When the Conflict Group is judged relative to the Control Group, the inverted V curve is further emphasized. Analysis of variance (see Table 2) reveals that the difference between the Conflict and Control Group in response to the stimulus dimension is significant at the .05 level. Analysis of variance of the Conflict Group alone yields significance (.05 level) for stimulus relevance, while analysis of the Control Group alone does not.

When all four groups are compared, the subjects who told a hostile story to a picture of low relevance are found to have significantly (.02 level) higher GSRs across the stimulus dimension than the subjects who did not. No other source of variance is significant. Inspection of the means reveals that, as with reaction time, the Uniform Inhibition Group produces a more exaggerated inverted V curve than the Conflict Group, as indicated by a greater decrease in GSR to the words of high relevance to hostility. The means, in micromhos, in order of increasing stimulus relevance for the two groups are: Uniform Inhibition Group: .66, .80, .58; Conflict Group: 1.06, 1.48, 1.36. The Uniform Expression Group and the Control Group show relatively little response variation as a function of the stimulus dimension, with means, in order of increasing stimulus relevance, respectively, of 1.07, 1.04, 1.13 and .86, .75, .72 micromhos.

TABLE 3

FORM OF GSR CURVE AS A FUNCTION OF EXPRESSION OF HOSTILITY ON PICTURES OF LOW AND HIGH RELEVANCE

Group	Form of curve			
	Monotonic	Inverted V	Miscellaneous	Total
Conflict	3	4	1	8
Control	2	0	6	8
Uniform expression	0	4	4	8
Uniform inhibition	1	2	5	8

Examination of individual curves makes it apparent that analysis of the means misrepresents the form of the GSR curves, as a cancelling out process takes place. In Table 3, it can be seen that seven of the eight curves of the Conflict Group as compared to one of the eight curves of the Control Group are either inverted V shaped or monotonic. Maximal discrimination between the two groups is obtained when both curves are considered indicative of conflict. Pooling all the subjects who produced a hostile response to a picture of low relevance, i.e., the Conflict and Uniform Expression Groups, and comparing them to the subjects who did not produce a hostile response to a picture of low relevance, i.e., the Control and Uniform Inhibition Groups, 8 out of 16 of the former and 2 out of 16 of the latter are found to produce an inverted V curve (see Table 3). This indicates a direct relationship between the inverted V curve and the production of a hostile response to the pictures of low relevance. A chi square of 5.24 is obtained, which is significant at the .05 level. The level of reliability remains the same when the subjects who produced a monotonic gradient are omitted from the analysis of the inverted V curves. The groups do not differ on monotonic gradients. A similar comparison for the picture of high hostility relevance fails to reveal significant differences. Four of the 16 subjects who told a hostile story and 6 of 16 subjects who failed to do so, produce inverted V curves. Three in each group produce monotonic gradients. It may be concluded that producing a hostile response to a picture of low hostility relevance is directly associated with producing an inverted V shaped curve of GSR, while performance on the picture

of high hostility relevance appears to be unrelated to the form of the GSR curve.

DISCUSSION

Gradients of GSR and Reaction Time

From the point of view of the theoretical model (Epstein & Fenz, 1962), it is encouraging to find that the subjects assumed to be in conflict from the content of their verbal responses produce distinctive curves of GSR and reaction time to a stimulus dimension. Moreover, relationships were demonstrated across projective techniques and across a considerable time interval, indicating that the conflict measured and the measures themselves are reasonably stable and general. The demonstration of reliable differences with relatively small groups suggests that the measures are adequate for individual prediction, a conclusion consistent with previous findings (Epstein & Fenz, 1962). However, the results are not without complications for the theoretical approach on which the study is based, as it had been anticipated that monotonic gradients would be produced rather than inverted V shaped curves.

What is the explanation for the inverted V shaped curves? We have found similar curves in a study of conflict associated with parachuting (Epstein & Fenz, 1962). While that study was concerned primarily with the performance of novice parachutists on a word association test containing a dimension of cues varying in relevance to parachuting, a few experienced jumpers were tested. All novice jumpers produced monotonic gradients, and all experienced jumpers produced inverted V shaped curves. We have since verified the association of the inverted V shaped curve with experience, and have learned that there is a progressive shifting of the apex of the curve toward the more remote end of the stimulus dimension with increasing experience. Moreover, the inverted V shaped curve occurs only on the day of a jump, when gradients are at their steepest. All this leads us to believe that the curve is an outcome of a learned inhibitory reaction which serves to keep the level of activation within homeostatic limits.

In order to obtain further information on the meaning of the inverted V shaped curve within the context of the present study, questionnaires on conflict over hostility which had been filled out in an earlier study (Saltz & Epstein, in press) were examined. It was found that the subjects who produced inverted V shaped curves obtained the highest scores on conflict over hostility, those who produced monotonic gradients were next, and those who produced other forms of curves were lowest, with the latter two groups differing little compared to their differences with the first group. Inspection of individual items suggests that the inverted V shaped curve is associated with excessive guilt over expression of hostility. Subjects who produce a monotonic gradient endorse items suggestive of moderate guilt. It will be interesting to follow up the two kinds of curves in studies of people who impulsively act out their hostility, such as psychopathic criminals and assaultive patients, and in clinical cases selected for intense feelings of hostility but who do not normally act on them.

Pictures of Low and High Relevance as Measures of Conflict

It was hypothesized that a conflict group, consisting of individuals who produced a hostile response to a picture of low relevance for hostility and failed to produce a hostile response to a picture of high relevance, would produce steeper gradients of GSR and reaction time than a control group, who behaved in an opposite manner. While the hypothesis is not supported for monotonic gradients, it is if inverted V shaped curves are substituted or if both curves are combined. Of further interest is the finding that the conflict curves of reaction time are associated primarily with performance on the picture of high relevance, and the conflict curves of GSR are associated primarily with performance on pictures of low relevance. A possible explanation is that GSR is less under conscious control than reaction time, and pictures of low relevance tap a deeper level of defense than pictures of high relevance. If this is true, it would support the assumption that gradients of inhibitory tendencies as a function of increasing cues are steeper than gradients of expressive tendencies which they inhibit, and that the latter can best be measured by pictures of

low relevance and the former by pictures of high relevance.

Of some incidental interest was the finding of a significant direct relationship between hostile responses to pictures of low relevance and magnitude of GSR across the stimulus dimension. As the relationship was unanticipated, and stands in isolation, it is in particular need of verification. If replicated, it may indicate that hostile responses to pictures of low relevance and magnitude of GSR both reflect a state of high activation where hostility is a symptom, rather than a cause.

SUMMARY

The present study was undertaken to test the internal consistency of a theoretical model from which the following three indices of approach-avoidance conflict are derived: (*a*) a relative increase in strength of approach responses to stimuli of low relevance and a relative decrease in strength of approach responses to stimuli of high relevance, (*b*) a positive gradient of physiological activation as a function of a stimulus dimension, and (*c*) a positive gradient of cognitive deficit as a function of a stimulus dimension.

Thirty-two subjects were selected according to their thematic responses on a specially constructed thematic apperception test which they took as part of a different study 6 months previously. The Conflict Group consisted of eight subjects who produced a hostile response to a picture of low relevance for hostility but failed to do so to a picture of high relevance; the Control Group consisted of eight subjects who responded in the opposite manner. A third group, the Uniform Expression Group, consisted of eight subjects who produced hostile responses to both kinds of pictures, while a fourth group, the Uniform Inhibition Group, was composed of eight subjects who produced no hostile responses. Six months after the thematic test, the groups were tested on a word association test with a built-in dimension of hostility. Measures of GSR, as an index of activation, and of reaction time, as an index of cognitive deficit, were obtained. The major findings and conclusions can be summarized as follows:

1. The Conflict Group produced a significantly different curve from the Control Group for both GSR and reaction time as a function of the stimulus dimension. The curve, which was inverted V shaped, was interpreted as indicative of the inhibition of anxiety, or of anxiety mediating responses, in response to cues of high relevance to the area of conflict. It was concluded that gradients of GSR and reaction time provide effective measures of conflict in individuals.

2. Conflict curves of reaction time were found to be associated primarily with performance on the picture of high hostility relevance, while conflict curves of GSR were associated primarily with performance on pictures of low relevance. It was suggested that reaction time to a stimulus dimension and verbal responses to cues of high relevance tap a higher level of awareness than GSR to a stimulus dimension and verbal responses to cues of low relevance.

3. A direct relationship was found between producing hostile responses to pictures of low relevance for hostility and overall magnitude of GSR, possibly indicating a common factor of generally high reactivity or conflict produced frustration, where hostility is a symptom rather than a cause.

REFERENCES

EPSTEIN, S., & FENZ, W. D. Theory and experiment on the measurement of approach-avoidance conflict. *J. abnorm. soc. Psychol.*, 1962, **64**, 97–112.

MALMO, R. B. Activation: A neuropsychological dimension. *Psychol. Rev.*, 1959, **66**, 367–386.

MILLER, N. E. Comments on theoretical models illustrated by the development of a theory of conflict behavior. *J. Pers.*, 1951, **20**, 82–100.

MURRAY, H. A. *Thematic Apperception Test manual.* Cambridge: Harvard Univer. Press, 1943.

SALTZ, G., & EPSTEIN, S. Thematic hostility and guilt responses as related to self-reported hostility, guilt, and conflict. *J. abnorm. soc. Psychol.*, in press.

(Received October 12, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 351–354

ANXIETY AND THE SPIRAL AFTEREFFECT TEST[1]

JAMES H. BRYAN[2] AND EDWARD LODER[3]

Pennsylvania State University

Past research on the Spiral Aftereffect Test (SAET) has been devoted primarily to determining its efficiency as a diagnostic instrument. Studies by Blau and Schaffer (1960), Davids, Goldenberg, and Laufer (1957), Gallese (1956), London and Bryan (1960), Page, Rakita, Kaplan, and Smith (1957), Price and Deabler (1955), and Sinberg (1961) have indicated the potential value of this test for the detection of organic brain damage. Although there is evidence to the contrary (Berger, Everson, Rutledge, & Koskoff, 1958; Gilberstadt, Schien, & Rosen, 1958; and Goldberg & Smith, 1958), it has generally been found that this test successfully discriminates organically impaired subjects from control subjects who are not brain damaged.

There is, however, a notable dearth of studies designed to test various theoretical notions advanced to account for the empirical findings. London and Bryan (1960) have advanced an explanation of the discriminatory powers of the SAET and have presented evidence in its support. They have maintained that the failure of the organic subject to report the Spiral Aftereffect (SAE) is, at least in part, a function of the organic's lack of abstract ability and the consequent anxiety produced by the unusual nature of the stimulus conditions. That is, London and Bryan have argued that attitudinal factors play a primary role in the perception and/or reporting of the SAE. This study demonstrated that, under conditions designed to reduce anxiety, the organic will report the SAE as frequently as normals, will report the SAE more frequently than organic subjects not given anxiety-reducing instructions, and will report the SAE significantly more frequently when given anxiety-reducing instructions than when not given such instructions.

It was the purpose of this study to investigate further the hypothesis that anxiety might be a crucial determinant in the reporting of the SAE. Since London and Bryan assumed that anxiety was already present in the subjects at the time of testing, it was decided that further support would be given their theoretical position if anxiety was induced in a sample differing markedly from the original. Furthermore, since some question remains concerning the relationship between intelligence and the reporting of the SAE (Aaronson, 1958; Goldberg & Smith, 1958; Holland, 1957; Page et al., 1957), this relationship was also investigated.

The hypotheses tested were:

1. Reports of the SAE are significantly fewer when the SAET is administered under conditions designed to induce anxiety than when presented under conditions not so designed.

2. There is a significant inverse relationship between reports of the SAE and scores obtained on the Test Anxiety Scale for Children (TASC).

3. There is a significant inverse relationship between reports of the SAE and scores obtained on the General Anxiety Scale for Children (GASC).

METHOD

Subjects

The subjects were 96 randomly chosen fifth-grade children, whose mean age was 11.6 years with a range of 11.1 to 13.2 years. Sexes were equally represented in the sample. Each subject was randomly assigned to one of the two treatment conditions (anxious or nonanxious). The sole limitation upon the random assignment of the subjects was that each group should contain an equal number of males and females.

[1] Based on a thesis submitted by the junior author in partial fulfillment for the master of science degree in the Department of Psychology, the Pennsylvania State University, University Park, Pennsylvania. The authors wish to thank Ellen Piers and Helen Snyder for their assistance in this study.

[2] Now at the Neuropsychiatric Institute, UCLA School of Medicine, Los Angeles, California.

[3] Now at Allentown Public School District, Allentown, Pennsylvania.

Apparatus

The standard 920-degree Archimedes spiral was imprinted on a cardboard disc 7½ inches in diameter. The apparatus was battery operated and the direction of rotation could be manipulated through the use of a toggle switch. The rotation speed of the spiral was approximately 78 rotations per minute. The apparatus herein described is identical to that employed by Blau and Schaffer (1960).

Procedure

Exactly 3 weeks prior to the administration of the SAET, all subjects were given the General Anxiety Scale for Children, the Goodenough Draw-a-Man Test, and the Test Anxiety Scale for Children. In addition, the subject's score on the Otis Quick Scoring Mental Ability Test, Form C-M was available.

Immediately prior to the administration of the SAET, each subject was individually given some block design tasks. The experimental group (anxious group) was given the three most complex block designs of the Wechsler Intelligence Scale for Children (WISC), while the control group (nonanxious group) was given the three least complex designs of the WISC. During the administration of the block design tasks and the SAET to the experimental group, the following instructions were given:

> Sit there. This is an important test of the way you think and how quickly you can learn things, so don't make any mistakes. These blocks have different sides. They can be put together easily to make a design like this one on this card (the experimenter shows the subject Design V). Go ahead and make one exactly like it as quickly as you can. Remember you are being scored for time on this too, so work very quickly.

Before the subject had completed the design, the experimenter commented,

> Time's up. I thought everyone would get the first one. Do better on this one (the experimenter shows the subject Design VI). You should be able to get this one in plenty of time. Can't you go any faster than that? Well, I guess you are not going to get that one either (and the experimenter interrupts the subject's construction of Design VI). Here's another one (the experimenter shows Design VII). You probably won't even get this one, the way you have been failing. (Again the subject was not allowed to finish the task.) Time is up. You certainly didn't do very good. I'm surprised. Maybe I better give you our eye test now. Sit there so you can see the little silver button in the center of the wheel (the experimenter points to the acorn nut). I will ask you questions about what you see.

After having received the instructions for the introduction of the block design test, as designated by the WISC manual, the control group received the following instructions:

> Now make one just like mine. (After the subject completed the task) Good. Now see if you can make one like this. (During the subject's construction), take all the time you want for these; you are doing very nicely. (After the subject had completed Block Design II), these don't get much harder, but let's just do this next one. This time we're going to put the blocks together to make them look like this picture. Watch me first. You see, the top of these blocks look the same as this picture. Now you make one like it. (After the subject completed the task), fine. You did that very nicely. Now I'm going to have you take a short test to check your eyes, the way you see things. I'll use this machine, so please make yourself comfortable so that you can see the little silver button in the center of this wheel (the subject points to acorn nut). The box contains a motor that will turn the wheel. I will ask you questions about what you see.

Eight trials, four in which the spiral was rotated in a clockwise direction (A) and four in which the spiral was rotated in a counterclockwise direction (B), were given each subject. For half the subjects in each group, the order of rotation was ABBABBAA, while for the remaining half of the subjects, the rotation order was BAABAABB. Responses were scored 1 if correct and 0 if not. The criteria of an acceptable response was that the subject reported the aftereffect to be the reverse, or the opposite, of the correct effect he reported during rotation. Acceptable responses included responses which indicated that the spiral was expanding (or contracting) or the correct direction of the black line during the illusion (going away or coming towards). It was then possible for the subject to obtain a score which ranged from 0 to 8.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF EXPERIMENTAL AND CONTROL GROUPS ON PRELIMINARY MEASURES

	Measures							
	GASC		TASC		Goodenough IQ		Otis IQ	
Condition	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Anxious	19.29	7.69	14.48	5.78	93.04	13.42	105.11	13.55
Nonanxious	18.69	7.44	16.25	6.13	91.17	12.76	100.76	11.48

RESULTS

Table 1 indicates the means and standard deviations of the anxious and nonanxious groups on the preliminary measures. A series of t tests indicated that there were no differences between the groups on these measures.

Table 2 presents the means and standard deviations for scores obtained on the SAET.

Since the distribution of scores on the SAET was highly skewed, the Mann-Whitney U test was employed in the comparison of the two groups (Siegel, 1956). Comparison across groups indicated that subjects in the anxious group reported the SAE significantly less often than did nonanxious subjects ($z = 1.71$, $p < .05$).

Since other investigators (Sarason, Davidson, Lighthall, Waite, & Ruebush, 1960), as well as the present authors, have noted marked sex differences on both the TASC and the GASC, correlations between reports of the SAE and scores on the anxiety scales were computed by transforming scores on the latter scales into z scores. No correlation was found to be significantly different from zero; nor were there significant differences between the correlations.

TABLE 2

MEANS AND STANDARD DEVIATIONS FOR SAE SCORES OF THE EXPERIMENTAL AND CONTROL GROUPS

Group	M	SD
Anxious	6.79	2.02
Nonanxious	7.56	0.92

DISCUSSION

The results of this study support, albeit equivocally, the hypothesis that anxiety is one determinant in the reporting of the SAE. Although situationally induced anxiety did, in the present study, depress reporting of the SAE, neither the GASC or TASC scores showed the predicted relationship to SAE reports. The question then arises from both the Bryan and London study (1960) and the present study as to the nature of the anxiety which inhibits SAE reports. While these studies indicate that anxiety is a potent factor in the reporting of the SAE, studies by Page et al. (1957), Price and Deabler (1955), and Spivack and Levine (1957), found that subjects suffering from various emotional disorders, and presumably considerable anxiety, report the SAE more readily than do organically damaged subjects. Insofar as the GASC measures states of anxiety that may be similar in nature to that experienced by psychiatrically disturbed subjects, the results of the present study support those studies which failed to show the relationship between anxiety and SAE reports. It is tentatively suggested therefore that for anxiety to inhibit SAE reports, it must be specific to the situation; that is, it must be situationally induced. Whether this anxiety differs in nature or intensity from anxiety suffered by those diagnosed as emotionally disordered is still a moot point. Although London and Bryan (1960) have suggested that the anxiety experienced by the organic subject is primarily the result of the unusual stimulus qualities of the SAE, it is evident from the present study that, at least for children who are not suspected of suffering from organic impairments, other types of stimulus conditions giving rise to anxiety may also inhibit SAE reports. Whether this is also true of adults who do suffer such impairments is still an unanswered question.

Although several speculations are appropriate in explaining the failure of the SAE and the TASC to be significantly related, perhaps the most parsimonious is one revolving around the instructions to the subjects regarding the SAE. As Sarason et al. (1960) have indicated, the validity of the TASC is greatly influenced by the testing situation and the test instructions. Sarason and his colleagues have argued that the TASC will show the greatest predictive power in those situations in which the subject perceives he is to be evaluated by an authority figure and in which the situation approximates the school atmosphere. It is likely that, by structuring the SAET as an eye test, test anxiety as measured by the TASC, would be considerably reduced. That is, anxiety concerning quality of performance would not be particularly evident. Under such circumstances then, it would not be surprising to find no significant relationship between test anxiety and the SAE reports.

It should be emphasized that the above comments concerning the influence of anxiety

upon the reporting of the SAE are not intended to imply that the behavior of organically brain damaged subjects in response to the SAET is unrelated to tissue damage; the relevance of these comments lies in the implication that the failure of organics to report this effect need not necessarily be linked, in a direct fashion, to tissue damage.

It should be further noted that neither the results of the London and Bryan (1960) study nor those of the present study in any way reflect upon the inadequacy of the SAET as a diagnostic tool but rather, do point to the need of carefully attending to those stimulus conditions which may affect anxiety levels of the subject during the testing session. Certainly the potent effects of situational variables upon this test are not unique to this test only (cf. Masling, 1960).

The results of the present study add further confirmation to the accumulating evidence (Goldberg & Smith, 1958; Holland, 1957; Page et al., 1957) that there is no relationship between reports of the SAE and intelligence.

SUMMARY

This study tested the hypothesis advanced by London and Bryan (1960) that anxious subjects would report the SAE significantly less often than nonanxious subjects. The relationship of intelligence with reports of the SAE was also investigated.

Ninety-six, randomly selected, fifth-grade school children were randomly assigned to one of two treatment conditions. The 48 subjects comprising the control group (nonanxious group) were individually given the three least complex block design tasks of the WISC. Throughout this period, they were allowed to finish each design, and received both encouragement and praise for their performance. The remaining 48 subjects (anxious group) were given the three most complex block design tasks of the WISC. In contrast to the control group, they were not allowed to finish the designs, and were given instructions designed to invoke anxiety. Immediately following this procedure, the subjects were given the SAET. Scores on both the TASC and the GASC as well as the Otis Quick Scoring Mental Ability Test and the Goodenough Draw-a-Man Test, were available.

Those subjects tested under conditions designed to invoke anxiety reported significantly fewer SAEs than did the control subjects. No significant relationship was found between the SAE reports and anxiety as measured by the TASC or the GASC, or between SAE reports and IQ measures.

REFERENCES

AARONSON, B. S. Age, intelligence, aphasia and the spiral aftereffect in an epileptic population. *J. clin. Psychol.*, 1958, **14**, 18–21.

BERGER, D., EVERSON, R., RUTLEDGE, L., & KOSKOFF, Y. D. The spiral aftereffect in a neurological setting. *J. consult. Psychol.*, 1958, **22**, 249–255.

BLAU, T. H., & SCHAFFER, R. E. The spiral aftereffect test as a predictor of normal and abnormal electroencephalographic records in children. *J. consult. Psychol.*, 1960, **24**, 35–42.

DAVIDS, A., GOLDENBERG, L., & LAUFER, M. W. The relation of the Archimedes spiral aftereffect and the trial-making test to brain damage in children. *J. consult. Psychol.*, 1957, **21**, 429–433.

GALLESE, A. J., JR. Spiral aftereffect as a test of organic brain damage. *J. clin. Psychol.*, 1956, **12**, 254–258.

GILBERSTADT, H., SCHEIN, J. D., & ROSEN, A. Further evaluation of the Archimedes spiral aftereffect. *J. consult. Psychol.*, 1958, **22**, 243–248.

GOLDBERG, L. R., & SMITH, P. A. The clinical usefulness of the Archimedes spiral in the diagnosis of organic brain damage. *J. consult. Psychol.*, 1958, **22**, 153–157.

HOLLAND, H. Drugs and personality: The effects of stimulus on visual aftereffects. *J. ment. Sci.*, 1957, **103**, 650–655.

LONDON, P., & BRYAN, J. H. Theory and research on the clinical use of the Archimedes spiral. *J. gen. Psychol.*, 1960, **62**, 113–125.

MASLING, J. The influence of situational and interpersonal variables in projective testing. *Psychol. Bull.*, 1960, **57**, 65–85.

PAGE, H. A., RAKITA, G., KAPLAN, H. D., & SMITH, N. B. Another application of the spiral aftereffect in the determination of brain damage. *J. consult. Psychol.*, 1957, **21**, 89–91.

PRICE, A. C., & DEABLER, H. L. Diagnosis of organicity by means of the spiral aftereffect. *J. consult. Psychol.*, 1955, **19**, 299–302.

SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., WAITE, R. R., & RUEBUSH, B. K. *Anxiety in elementary school children*. New York: Wiley, 1960.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill, 1956.

SINBERG, R. M. Some effects of stimulus variation on spiral aftereffect in organic and nonorganic subjects. *J. consult. Psychol.*, 1961, **25**, 129–136.

SPIVAK, G., & LEVINE, M. The spiral aftereffect and reversible figures as measures of brain damage and memory. *J. Pers.*, 1957, **25**, 767–778.

(Received October 23, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 355-361

PERCEPTUAL AND COGNITIVE FUNCTIONING OF RETARDED READERS

RICHARD H. WALTERS AND HELEN DOAN [1]

University of Toronto

In a study designed primarily to examine personality difficulties of retarded readers with particular reference to psychoanalytic theory, Walters, Van Loan, and Crofts (1961) noted that these readers performed more poorly on selected perceptual and perceptual-motor tasks than did average or advanced readers. In one of the tasks the children were required to discover in which of four compartments of a multiple-choice apparatus a nude doll had been hidden. Retarded readers were slower to open the compartment in which they thought the doll was hidden than were either of the other two groups, a result which could be considered to support the psychoanalytic theory that one characteristic of the retarded reader is a fear of looking based on early sexual experiences. However, a more detailed analysis of the data suggested that similar differences among the three groups of children might have been obtained if a non-sexual test object had been employed. The suggestion was supported by the finding that retarded readers also performed relatively poorly on perceptual closure tasks which seemed devoid of sexual significance. Consequently, it was decided to compare the performance of advanced, average, and retarded readers on a further series of perceptual, perceptual motor, and cognitive tasks, this time without reference to psychoanalytic theory.

The activity of reading appears to require at least two important perceptual functions—the recognition of the form of words and the ability to distinguish small details of shape (cf. Vernon, 1958). In addition, it requires the association of a symbol with an object or event, a process that may be regarded as cognitive rather than perceptual. On the basis of the above considerations, it was predicted that retarded readers would perform more poorly than average or advanced readers on each of the following tasks: The Steer-Beatty Closure-Threshold Test, used in the study by Walters et al. (1961); the Perceptual Differentiation Test employed by Gibson and Gibson (1955) to study the development of perception in children and adults; and a learning task involving the association of visual symbols with objects. In addition, the multiple-choice task used by Walters et al. (1961) was repeated, this time with a presumably neutral "nebbish" as the test object instead of the nude male doll used in the earlier study. It was expected that, in general, advanced readers would give the best performance on all tasks, with average readers performing at a level intermediate between that of the advanced and retarded readers.

Lachman (1960) has pointed out that the hypothesis favored by Bender (1949) and others that retarded readers suffer from a lag in perceptual development overlooks the role that emotional factors play in reading disability. Walters et al. (1961) suggested that these emotional factors may be related to parental frustration and punishment of exploratory behavior, which results in an inhibition of curiosity (the psychoanalysts' "fear of looking") and poor achievement motivation. If this is correct, the promise of a reward should have more influence on the performance of retarded readers on relevant perceptual and cognitive tasks than on that of readers who are performing at an adequate level. This hypothesis concerning the provision of an incentive assumes both that reading retardation is a symptom of a generalized perceptual and cognitive lag and that emotional (or motivational) factors are involved in reading retardation. It was tested through the use of the 3×2 factorial design described below.

[1] The authors are indebted to the Scarborough Township Board of Education and to the staff of the Corvette and Knob Hill Schools for their cooperation in this study. Marian Hyatt Jerry assisted in the collection of data.

METHOD

Subjects

The subjects were 54 Grade 7 and Grade 8 boys from two suburban public schools in the same area of Metropolitan Toronto. They were selected from the total population in these grades after the administration of a group intelligence test (The Dominion Test of Learning Capacity, Form A—Intermediate) and a test of reading ability (The Dominon Group Achievement Test, Niagara Edition: Diagnostic Paragraph Comprehension—Grade 8) which have been standardized on Canadian children. Standard scores were calculated on each test for each of the four groups of children separately. Each child's z score on the intelligence test was then subtracted from his z score on the reading test. It was decided to regard a child with a discrepancy score of $+5$ as an advanced reader, a child with a discrepancy score of -5 as a retarded reader, and a child whose discrepancy score lay between $+2$ and -2 as an average reader. Because of the difficulty of assessing the meaning of reading retardation in bright and dull children, children whose IQs were above 120 or below 90 were eliminated from the possible pool of subjects.[2] In addition, those children whose medical histories showed records of visual or auditory anomalies or behavior problems, or who came from broken homes, were excluded from the study. From the remaining children, 27 were selected from each school in such a way as to maximize the size of discrepancy scores of advanced and retarded readers and to minimize the size of the discrepancy scores of average readers.

The subjects were equally divided among the schools in order to study the effects of reward on the performance of the three groups of children. If some children receive rewards for participation in a project and others do not, and the children are aware of this distinction, resentment is bound to follow and influence in unknown ways the test results. Consequently, children in School A were assigned to the rewarded groups and children in School B to the nonrewarded groups. In addition, it was decided to balance as well as possible the number of Grade 7 and Grade 8 boys in each of the three reading groups. As a result of these restrictions and those described in the previous paragraph, the number of available children was reduced to nine in each of the six experimental groups.

Apparatus

The equipment required for two of the four tests has been described in detail in a previous publication (Walters et al., 1961). The Steer-Beatty Closure-Threshold Test consisted of 12 sets of 5 cards, each card in a series being a representation of a single object. Because of the relative spacing of dots in the figure and dots in the background, the figure becomes increasingly easier to identify as the subject progresses from the first to the last card in a series. The multiple-choice apparatus was a box separated into four compartments. When the back of the box and the front doors of the compartments were closed, a buzzer was sounded and a timer was set into operation. When the experimenter held the back of the box open or the subject made his choice by opening a compartment door, the circuit activating the buzzer and timer was broken.

The perceptual differentiation test was that used by Gibson and Gibson (1955) to study the development of form perception in children and adults. It consisted of a pack of 2 inches $\times$ 4 inches cards of nonsense "scribbles." The pack of cards contained a standard, 4 cards identical with the standard, 17 cards that differed from the standard in not readily identified ways, and 12 cards that were readily distinguishable from the standard. The subject's task was to select the four cards that perfectly matched the standard.

For the symbolic learning task a plywood board was fitted above the previously described multiple-choice apparatus (Figure 1). Three rows of four lights, 5 inches apart horizontally and 2½ inches apart vertically, were mounted on the board. The lights were of four different colors—red, yellow, green, and blue—randomly displayed on the board. When either the rear or one of the front doors was open, the panel of lights was disconnected. When the box was completely closed, one of the lights was illuminated at the same time as the buzzer sounded and the timer started. A panel of buttons at the back of the box was used to select the light to be illuminated on any one trial.

Procedure

The subject was conducted by the experimenter to a well-illuminated room with two decks side by side, on one of which stood the multiple-choice box. The subject was instructed to sit in front of the empty desk, facing the experimenter, who seated herself at the opposite side.

The Steer-Beatty cards were first presented with the following instructions: "I am going to show you some pictures in each of which a thing is hidden. I want you to tell me what that hidden thing is. For example, can you tell me what is hidden in this picture?" (For purposes of demonstration the subject was given two cards from the Steer-Allen Figure-Ground Confusion Test, Walters et al., 1961, one representing a plate and another a knife.) When it was evident that the subject had understood the nature of the task, the experimenter continued: "The rest of these will be a little more difficult. Take your

[2] The mean IQs for advanced, average, and retarded readers were 106.28, 103.50, and 105.67, respectively. Since reading retardation reduces scores on group IQ tests which include verbal items, the retarded readers were undoubtedly, in some senses, "brighter" than the other two groups of children. This circumstance probably reduces the chances of obtaining significant differences in the predicted direction on many nonverbal tasks.

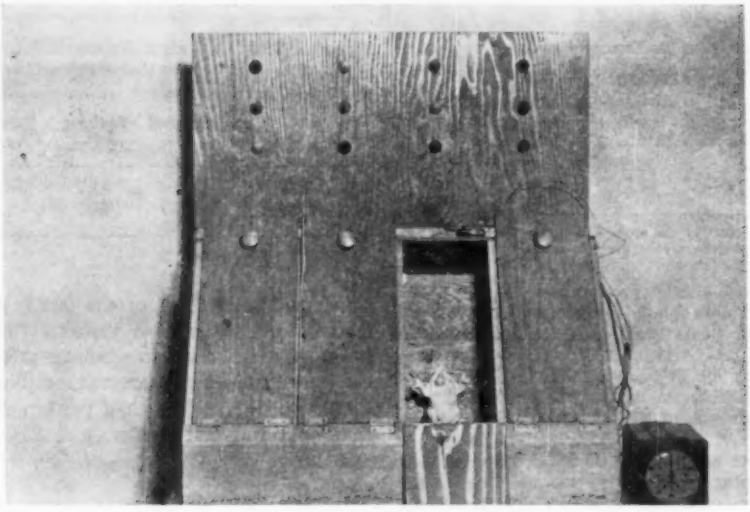

FIG. 1. Multiple-choice apparatus used for the symbolic learning task.

time and tell me what is hidden in each of these pictures." The Steer-Beatty cards were now presented.

There was no time limit to each response in the test. If the subject could not recognize the object, the experimenter presented the next card. The total number of trials necessary to recognize all the objects was recorded by the experimenter.

The subject was now asked to stand in front of the discrimination box. The experimenter then stated:

> I am going to place this nebbish (holding up the test object) in one of these four boxes. I want you to see if you can find it. You can do this by opening one of these doors as soon as you hear the buzzer. (Demonstrated) When you hear the buzzer you can open the door in which you think the nebbish is hidden. Don't close it until I tell you.

The nebbish was placed into the back of the various compartments in a predetermined random order. The experiment continued until the subject had completed 40 trials. The latency of the subject's response and his choice of box were recorded for each trial.

The subject was then asked to sit down once more in front of the other desk. The following instructions were given: "I am going to show you a card, on the front of which a pattern is printed. Please examine it carefully." (The standard item of the Gibson and Gibson Perceptual Differentiation Test was then shown to the subject for 5 seconds.) The experimenter continued: "There are some items in this pack of cards exactly like the one I showed you." The series of items, including the four identical copies of the standard, was then presented by the experimenter one at a time for 3 seconds each. The subject was not told whether his identification was correct or incorrect. At the end of the series, if the subject had failed to extract the four replicas of the standard, he was again shown the standard for 5 seconds, and the cards were again presented, each time in a different random order. The experimenter scored the total number of trials necessary to identify the four critical cards on a single trial.

The subject was now told to stand in front of the discrimination box. The instructions were as follows:

> Now, I am going once more to hide the nebbish in one of the compartments of the box. This time, I want you to watch the lights. If you watch carefully, you can learn to choose the correct box every time. As soon as you see one of the lights go on, open the door of the box in which you think the nebbish is hidden.

The procedure was continued until the subject learned to associate one light with each compartment for a total of eight trials. The compartments, from left to right, were symbolized by blue, green, yellow, and red in that order. After the criterion of learning had been reached, the color-compartment associations were reversed, so that blue now stood for the right-hand compartment and red for the left. Trials were continued until the subject had made eight successive correct responses.

Finally, each subject was presented with a color naming chart and instructed to name the colors as quickly as possible working from left to right along each line. This was continued until it was obvious that the subject had no difficulty in identifying colors.

TABLE 1

MEAN SCORES ON THE PERCEPTUAL CLOSURE TASK

Group	Advanced readers	Average readers	Retarded readers
Rewarded	37.00	30.67	32.00
Not rewarded	34.89	36.22	37.33

Note.—The smaller the score, the better the performance. $N = 9$ in each group.

The subjects in School A had been assigned to the reward condition. Before commencing Task 1, these subjects were told: "If you perform well enough on the various tasks which I will ask you to do, you can have any one of these prizes." (The subject was shown some miniature sets of checkers, metal puzzles, toy automobiles, etc.) "Remember to do the best you can."

In order to minimize the effects of communication between subjects, each subject was asked not to tell anyone about the tasks until the experiment was finished and, as far as possible, in each of the schools the order of testing the subjects was balanced over the three groups of readers. In addition, the subjects in the reward condition were given only a small plastic chip at the end of the experimental session; these chips were exchanged for toys only when all subjects in School A had been tested. This latter procedure served to lessen the attention given to subjects at the end of the testing period and to reduce the possibility of their being questioned by subjects who had not yet been tested.

RESULTS

Perceptual Closure

Table 1 gives the mean scores for the three groups of readers under the reward and nonreward conditions. None of the observed differences reached an acceptable level of significance (Table 2).

Perceptual Differentiation

The mean number of trials required by each of the six groups of subjects to differentiate the four critical cards of the test is given in Table 3. An analysis of variance (Table 4) indicated that differences among groups and the interaction effect were not significant, but that rewarded subjects had performed better than nonrewarded subjects ($p < .05$).

TABLE 2

ANALYSIS OF VARIANCE OF PERCEPTUAL CLOSURE SCORES

Source	*SS*	*df*	*MS*	*F*
Groups	56.259	2	28.130	<1
Conditions	115.574	1	115.574	2.79
Groups × Conditions	171.371	2	85.686	2.07
Error	1,988.444	48	41.426	

TABLE 3

MEAN NUMBER OF TRIALS REQUIRED FOR CORRECT PERFORMANCE ON THE DIFFERENTIATION TEST

Group	Advanced readers	Average readers	Retarded readers
Rewarded	2.56	4.22	4.44
Not rewarded	6.00	4.00	6.11

Note.—$N = 9$ in each group.

Reaction Times

Median latencies were computed for each subject for each block of 10 consecutive trials (Trials 1 to 10, 11 to 20, etc.). The means of these latencies for the three groups of subjects are plotted in Figure 2. A Type III analysis of variance (Lindquist, 1953) yielded a significant Groups × Conditions × Trials effect (Table 5). As Figure 2 indicates, advanced readers gave the slowest reaction times. The provision of reward seemed to influence the performance of both advanced and retarded readers, its effect being most marked at the commencement of testing.

Symbolic Learning

The mean number of trials required by each group of subjects to associate a light of a particular color with the appropriate compartment is given in Table 6. An analysis of variance (Table 7) yielded a significant *F* value

TABLE 4

ANALYSIS OF VARIANCE OF PERCEPTUAL DIFFERENTIATION SCORES

Source	*SS*	*df*	*MS*	*F*
Groups	14.333	2	7.167	1.03
Conditions	35.852	1	35.852	5.14*
Interaction	30.259	2	15.130	2.17
Within	334.889	48	6.977	

* $p < .05$.

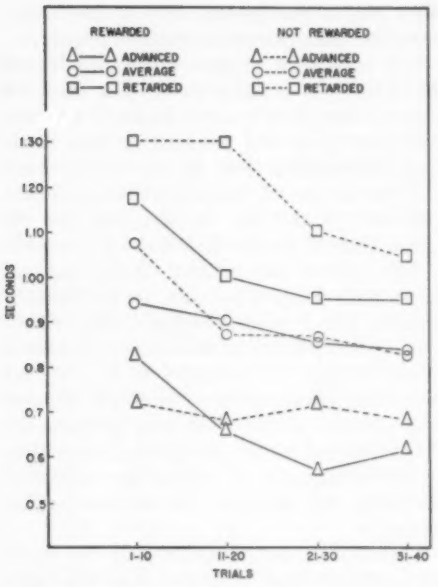

FIG. 2. Mean reaction times of advanced, average, and retarded readers under reward and nonrewarded conditions.

for differences among groups ($p < .005$). Retarded readers performed more poorly than either advanced or average readers, the difference between the latter two groups being negligible.

The mean number of trials for each group to learn the reversal task is given in Table 8.

TABLE 5

ANALYSIS OF VARIANCE OF MEDIAN SCORES ON THE REACTION TIME TASK

Source	*SS*	*df*	*MS*	*F*
Groups	7.518	2	3.759	3.76*
Conditions	0.408	1	0.408	<1
Groups × Conditions	0.339	2	0.170	<1
Error	48.028	48	1.001	
Trials	1.581	3	0.527	11.71***
Trials × Groups	0.117	6	0.019	<1
Trials × Conditions	0.048	3	0.016	<1
Trials × Groups × Conditions	0.685	6	0.114	2.53*
Error (w)	6.514	144	0.045	

* $p < .05$.
*** $p < .001$.

TABLE 6

MEAN NUMBER OF TRIALS TO CRITERION ON THE SYMBOLIC LEARNING TASK

Group	Advanced readers	Average readers	Retarded readers
Rewarded	53.89	39.78	71.89
Not rewarded	32.67	31.11	128.67

Note.—$N = 9$ in each group.

An analysis of variance (Table 9) showed that nonrewarded subjects had performed more poorly than rewarded subjects. The interaction effect approached, but did not quite reach, an acceptable level of significance ($p < .10 > .05$).

TABLE 7

ANALYSIS OF VARIANCE OF NUMBER OF TRIALS TO CRITERION ON THE SYMBOLIC LEARNING TASK

Variance	*SS*	*df*	*MS*	*F*
Groups	45,082.333	2	22,541.167	6.46**
Conditions	1,084.519	1	1,084.519	<1
Groups × Conditions	15,786.926	2	7,893.463	2.26
Within	188,410.222	48	3,489.078	

** $p < .005$.

DISCUSSION

Relatively few investigations of reading disability have been concerned with the perceptual and symbolic learning of retarded readers. This study points to the importance of further research in this area. The findings are by no means clear-cut; however, the subjects were all within a range of reading ability that permitted them to receive the usual grade-school instruction. While the inclusion of children with reading disabilities sufficiently great

TABLE 8

MEAN NUMBER OF TRIALS TO CRITERION ON THE REVERSED SYMBOLIC LEARNING TASK

Group	Advanced readers	Average readers	Retarded readers
Rewarded	6.89	4.00	4.67
Not rewarded	7.11	8.11	21.67

Note.—$N = 9$ in each group.

TABLE 9

ANALYSIS OF VARIANCE OF NUMBER OF TRIALS TO CRITERION ON THE REVERSED SYMBOLIC LEARNING TASK

Source	*SS*	*df*	*MS*	*F*
Groups	536.926	2	268.463	2.07
Conditions	682.667	1	682.667	5.27*
Groups × Conditions	694.110	2	347.055	2.68
Within	43.623	48	0.909	

* $p < .05$.

to demand assignment to special classes or attendance at a reading clinic might have brought more conclusive results, it would have complicated the task of interpretation. Special treatment of school children often creates emotional problems over and above those involved in the disability it is intended to correct.

Although differences were not significant, certain trends on the perceptual closure task may be noted. The results for nonrewarded subjects are comparable to the finding of Walters et al. (1961) that the performance of both advanced and average readers surpassed that of retarded readers. Under reward conditions, however, retarded readers gave as good a performance as average readers and a better one than advanced readers. The prediction that the provision of reward would generally improve performance was not supported, since the performance of both the average and advanced readers who were rewarded was poorer than that of the comparable nonrewarded groups.

The results from the perceptual differentiation task yielded a trend which agrees with the finding of Coleman (1953) that retarded readers do not perform well on tasks requiring perceptual discrimination. The results of the two perceptual tasks, i.e., the perceptual closure and differentiation tasks, were somewhat similar. Average achievers, on the whole, gave the best performance. Underachievers, on the other hand, gave the worst performance. The fact that average achievers performed as well as overachievers on these two tasks may be due to inadequate selection procedures. On the other hand, a *more than average* level of reading skill may be due to factors other than perceptual-motor maturity.

The advanced readers performed better throughout the multiple-choice task than did either of the other groups. Figure 2 suggests that each group was reaching its own maximum performance level by the last 10 trials and that the rate of improvement was initially increased, in the case of advanced and retarded readers, by the provision of a reward. A high level of performance on this task requires both a rapid response to an auditory stimulus and a rapid selection of the "right" compartment. Raab, Deutsch, and Freedman (1960) report that retarded readers are no slower in responding to an auditory stimulus than are good readers. The poor performance of the retarded readers in this study may thus be due to cognitive or personality difficulties (or both) that influence the decision making process.

The most important findings of this study are probably those obtained from the symbolic learning task. They suggest that the difficulty that retarded readers have in responding to verbal symbols is one aspect only of a more general inability to link symbols with objects or actions. Since the colored lights were extremely simple visual stimuli, the poor performance of retarded readers cannot be attributed to a delay or arrest of perceptual learning of the kind required for successful performance of the differentiation task. It would appear, rather, that an important higher-order cognitive function is only weakly developed in these readers. Their responses to nonvisual symbols will receive attention in a future study.

To overcome practical difficulties, children in School A were assigned to the reward condition and children in School B to the nonreward condition. Thus main effects attributed to reward could be partly or wholly due to other factors associated with attendance at different schools. This is, however, unlikely, since all the children came from the same general area of the city, and from schools governed by the same educational policies. The home backgrounds of children attending both schools are quite diverse, with parental occupations ranging from the professional to the unskilled categories.

The provision of reward did not, in any case, have consistent effects on the performance of the three groups of subjects. On the perceptual differentiation and reaction time tasks, reward appeared to have some influence on the performance of both advanced and retarded readers. On the symbolic learning task it produced a quite marked improvement for the retarded readers. Although the Groups × Conditions interactions never quite reach an adequate significance level, there is some indication that the performance of retarded readers, particularly in the area of symbolic learning, can be improved by the provision of a reward. Overachievers may be children who are readily affected by incentives, even when rewards have been adequate in the past, while underachievers may be readily affected by incentives because past lack of incentives has produced modifiable habits of responding at less than maximum efficiency. A more conclusive indication that this is the case might have been obtained if the attractiveness of rewards had been increased through the selection of more elaborate toys appropriate to the age level of each subject who was tested.

SUMMARY

Eighteen advanced, 18 average, and 18 retarded readers were given a series of perceptual and cognitive tasks. Half the subjects in each group were promised a reward for good performance.

In general, advanced and average readers performed better than retarded readers on the test battery; the latter were relatively weak in perceptual discrimination and symbolic learning and gave slower reaction times in a multiple-choice task.

The provision of a reward did not influence performance in a consistent manner; however, on most tasks it improved the performance of both advanced and retarded readers but had little effect on that of average readers.

Retarded readers appear to have difficulty in decision making and in associating a symbol with an object, i.e., difficulties at a cognitive level, as well as perceptual handicaps. In addition, they may lack the motivation to perform at maximum efficiency.

REFERENCES

BENDER, LAURETTA. Psychological principles of the visual-motor gestalt test. *Trans. NY Acad. Sci.*, 1949, **11**, 164–170.

COLEMAN, J. C. Perceptual retardation in reading disability cases. *J. educ. Psychol.*, 1953, **44**, 497–503.

GIBSON, J. J., & GIBSON, ELEANOR J. Perceptual learning: Differentiation or enrichment? *Psychol. Rev.*, 1955, **66**, 32–41.

LACHMAN, F. M. Perceptual-motor development in children retarded in reading ability. *J. consult. Psychol.*, 1960, **24**, 427–431.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.

RAAB, SHIRLEY, DEUTSCH, M., & FREEDMAN, A. M. Perceptual shifting and set in normal school children of different reading achievement levels. *Percept. mot. Skills*, 1960, **10**, 187–192.

WALTERS, R. H., VAN LOAN, MALLE, & CROFTS, IRENE. A study of reading disability. *J. consult. Psychol.*, 1961, **25**, 277–283.

VERNON, MAGDALEN D. *Backwardness in reading: A study of its nature and origin.* Cambridge, England: Cambridge Univer. Press, 1958.

(Received October 27, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 362–378

ZEN BUDDHISM:

A PSYCHOLOGICAL REVIEW

EDWARD W. MAUPIN

University of Michigan

Zen, a sect of Mahayana Buddhism, originated in China and has played an important role in Japanese culture since its introduction there in the thirteenth century. It has traditionally sought to bring about in its students a direct experience of the enlightenment which characterized the Buddha. What makes this of interest to psychologists is that enlightenment is considered a psychological problem to be worked out by the student. Appeals to divine intervention or intensive study of scriptures are felt to be irrelevant.

Zen involves a variety of training techniques designed to guide the student to a turning point, *satori*, which appears to be a major shift in the mode of experiencing oneself and the world, and which is an important step on the way to enlightenment. Since the individual, with satori, is described as living an increasingly effective and satisfying life, Zen is of interest in terms of psychotherapy. There is a growing body of literature by psychologists and psychiatrists on this aspect of Zen, and it seems likely that there are concepts and procedures here which will prove useful in Western psychotherapy. It is the purpose of this paper to distill out of the often confusing literature in English a coherent picture of Zen, its procedures and the experiences which result. Without wishing to overinterpret or "pigeon-hole" Zen experience, I have examined psychological literature which seemed to shed further light on the phenomena described.

After an overview of satori and the training process, we will proceed to interpretations which have been offered for satori and similar experiences, and then examine the relevance of the concept of regression in the service of the ego. Next, literature concerning the effects of waking consciousness on perception and action will be examined. Relaxation therapy and research related to attention appear to have links with the training process and will be taken up next. Finally, therapeutic application will be considered.

The Ordinary Adult Mind and Satori

Zen literature makes no particular distinction between types of psychopathology. As a therapy it seems designed for people who are normally mature and have achieved a fair degree of self-control. Existential problems are seen as resulting from the way the ordinary adult experiences himself and his world —from the terms in which the problems are couched. For an answer, a radical shift in the mode of experience, satori, is proposed.

Satori is not a trance. Consciousness is not lost, nor does the ability to use ordinary cognitive functions become impaired. It is not a quietistic retreat. All of these possible outcomes of the training are considered byways to be guarded against with the help of the Zen master. Satori is described as an added mode of experience, comparable to the opening of a third eye. It is considered impossible to express in rational language.

This raises the problem of irrationality in Zen literature, which warrants a short digression. The confusing, nonlogical quality seems to stem from three main sources. First there is the ordinary difficulty in describing any state of consciousness. Under the proper circumstances we can specify the content of consciousness, but the formal qualities are much more difficult to communicate. One recourse is to speak in analogies and hope that it "rings a bell" for the hearer. We find one Zen master counseling his students to keep a kind of "doubt" which arises in the course of meditation "neither too fine nor too coarse" (Luk, 1960). This first source of unclarity, then, is one which often plagues the attempt to communicate subjective experience.

A second source is the teaching method of Zen. The Zen master seeks to have the student get beyond concepts of satori to the experi-

ence itself. The student may come with a question about some aspect of Buddhism, the training, or his own problems in reaching satori. But a direct conceptual answer would only be *about* the topic; it would not bring the student to see the thing itself. There is a deep feeling in Zen that conceptual knowledge can come only so close to its object. So the Zen master may make an apparently illogical retort to jolt the student into seeing the thing for himself.

The third source of unclarity seems to be a genuinely alogical quality of satori itself. Certain aspects of this new mode of experience, such as the feeling of oneness, seem genuinely inexpressible in a language posited on a subject-object dichotomy, conventional time, space, and so on.

Although descriptions of satori are given with a caution that they are only inadequate analogies, there are certain uniformities in the way people compare it with the ordinary mode of experience. The first contrast is between intellection and intuition. Suzuki writes that man tends to mistake his conceptual tools for reality. "He forgets that concepts are his own creations and by no means exhaust reality. Zen is fully conscious of this and all its *mondo* are directed towards casting off the false mask of conceptualization" (1949a, p. 28).

Chang (1959, p. 141) outlines other characteristics of the ordinary mind. It must break reality up into discrete entities and can deal with only a few things at a time. It is rigid and fixed, unable to deal with all possible aspects of a thing, and it tends to "cling" to the object thus separated out and objectified. The term "clinging" hints at personal motivation to maintain the stability achieved by this kind of structuring. It will be noted that the conceptual mode is under attack not because it is useless in general, but because, improperly used, it cuts the individual off from another, more direct contact with his experience. It is this loss of immediate experience which plunges the individual into existential problems.

The tendency to live in a fantasy world (which treats objects and events in terms of a personal—and necessarily cognitive—network of goals, plans, wishes, and fears) is under attack. Also included is the tendency to impose too quickly the conventional structuring of time, space, the subject-object dichotomy, and self-other value systems. It should be noted, too, that the psychological unconscious, as repository of repressed derivatives of unacceptable wishes, is included in this conceptual filter. Suzuki is inclined to think of this as merely a lower stratum of the same kind of conceptual mind, "probably accumulated ever since we began to become conscious of our own existence" (1949a, p. 95).

Contrasted with this is "intuition," which Kondo, a Japanese psychoanalyst, interprets as "the function of the human mind for perceiving totality" (1952). Suzuki (Fromm, Suzuki, & DeMartino, 1960) speaks of it in terms of the more traditional *prajna:*

> Prajna is the experience a man has when he feels in its most fundamental sense the infinite totality of things, that is psychologically speaking, when the finite ego, breaking its hard crust, refers itself to the infinite which envelops everything that is finite and limited and therefore transitory. We may take this experience as being somewhat akin to a totalistic intuition of something that transcends all our particularized, specified experiences (p. 74).

Intimately connected with the conceptual mode of experience is the experience of self. The individual mistakes the self which he can take as an object of consciousness for his real self. Like other objects of consciousness it must be separated out, and this, in particular, leads the individual into a frantic scramble to defend, maintain, and bolster himself. Herrigel (1960) writes:

> By learning to discriminate himself more and more from everything that is not himself, that does not belong to him, man experiences the tension between the ego and non-ego as an opposition. The more consciously he confronts everything not himself as an object, the more the ego places itself outside—outside what is "opposite" to it. The result is a continuous division of being into the two realms of subjective and objective (p. 19).

It will be seen from this that the experience of a separate self is felt to stand in a causal relationship to other aspects of the clinging, conceptual mind. But from the standpoint of satori this self is a fiction. (Compare Sartre, 1957, p. 97.) Satori, in contrast, is the intuitive seeing into the real self, the true author

of one's behavior, which is at the same time a part of the whole flux of the universe. This contact, however, does not mean reflexive awareness of the real self—one can no more grasp it intellectually than before. Rather, one's experience is felt to take place directly through the real self unmediated by conscious thought, and without consciousness of the process. "It is rather a state of mind in which there is no specific consciousness of its own workings" (Suzuki, 1949b, p. 105). One is content to let behavior unfold out of a self which cannot be fully conceptualized. One trusts this self enough to suspend conscious reflective control over it.

Two passages (Suzuki, 1956) express better than any summary statement in what way satori is felt to resolve existential problems:

> Zen in its essence is the art of seeing into the nature of one's being, and it points the way from bondage to freedom. . . . We can say that Zen liberates all the energies properly and naturally stored in each of us, which are in ordinary circumstances cramped and distorted so that they find no adequate channel for activity. . . . It is the object of Zen, therefore, to save us from going crazy or being crippled. This is what I mean by freedom, giving free play to all the creative and benevolent impulses inherently lying in our hearts. (p. 33ff.)
>
> This very "moment" is not subject to birth-and-death and therefore there is no going beyond it as long as we live in this present moment. Here is absolute tranquility which is no other than this present moment. Bliss lies in the timelessness of this present moment. There is here no particular recipient of this bliss and, therefore, everyone of us is blissed with eternal bliss. (p. 111, Quoted from the ancient monk, Yeno.)

It is important to note that ordinary perception and conceptualization are not impaired in satori. Rather, while intuiting totality, the individual is said to perceive objects more objectively than before, less distorted by personal motives. Cognitive skills remain available as required.

Thus far we have discussed satori as though it were a constant state of mind after a certain point in Zen training. Actually, the first experience probably lasts only a short time. Further training is directed toward expanding the situations in which such a state of mind may be maintained. Enlightenment, a nearly impossible ideal, is apparently a constant experience of satori. Nevertheless, at times following satori, the individual may deepen the experience to a point of complete union. Relaxing conceptual consciousness and falling back on unreflected intuition is felt to have important effects on the range of stimuli to which one can respond. Zen-trained artists and soldiers especially have made use of this state of mind to heighten their effectiveness.

TRAINING

We turn now to the techniques which are used to lead the student toward the experience of satori. It must be emphasized that satori is felt to be an awakening to something which was always there, not the product of a technique. What procedures are used will depend on what the Zen master feels the student needs to awaken him. Specific methods vary far more than the present review can convey. Generally, however, the student spends a portion of the day sitting motionless and engaging in concentration. The object of concentration varies considerably and may be changed as the student progresses. The aim is to suspend the ordinary flow of thoughts without falling into a stupor. The achievement of undistracted attention is the first means of coming to grips with the egoistic, conceptual mode of experience.

The position of the body is of some importance. It must be relaxed and comfortable, yet not supine and likely to induce sleep. Ordinarily the full-lotus position is used. The eyes are kept partly open, again to avoid sleep or stupor.

Three objects of concentration deserve special mention: breathing, the *hua t'ou*, and the *koan*. Concentration on breathing seems to be the simplest of these and is frequently used to develop the ability to concentrate. "Hua t'ou" is a Chinese word which means "antethought." It describes the state of mind of one who approaches himself with the question, "Who is this who calls on the name of Buddha?" (Luk, 1960). As such, it appears to define an attitute of detached observation toward one's mental activity, keeping in mind its subjective origin. The hua t'ou overlaps with the koan, a statement which is impossible of rational comprehension, but understandable to the person who has ex-

perienced satori. Any of a number of traditional exchanges between masters and monks may be taken as a koan. The idea is not to run through the words themselves, but to penetrate to their meaning, the state of mind which they express. Not all Zen sects use the koan exercise. Where used, it seems to deepen the intellectual crisis preceding satori and to produce a more vivid satori experience. De Martino (Fromm et al., 1960) suggests that the koan serves to crystallize and focus the desperate personal need to find an answer. As he describes it, the attempt to solve the koan becomes almost a surrogate for the struggle to solve one's life.

In the wealth of subjective reports it is generally possible to distinguish two main stages in the course of training before satori. There is an initial phase in which concentration, difficult at first, eventually becomes more successful. Relaxation and a kind of pleasant "self-immersion" begin to follow. At this point internal distractions, often of an anxiety arousing kind, come to the fore. Herrigel (1956) indicates that the only way to render this disturbance inoperative is "to look at it equably and at last grow weary of looking" (p. 55).

Eventually a second phase begins (Herrigel, 1956). This is a state "in which nothing definite is thought, planned, striven for, desired or expected, which aims in no particular direction and yet knows itself capable alike of the possible and the impossible, so unswerving is its power . . ." (pp. 55–56). Concentration seems to be accompanied by a sense of calm stillness, of energy and vitality, and a feeling of invulnerability (Chang, 1959; Kondo, 1952, 1958). Sato (1959) and his students, who undertook a special "accelerated" course of training from a Zen master, reached this point after about five days. Both Chang and Herrigel suggest that the phase begins with a "jolt" or "shock," but this is not invariably mentioned.

This state of mind is traditionally described with the analogy of a mirror, which reflects many things, yet is itself unchanged by them. It seems likely that this phase of meditation, in particular, increases receptivity to previously excluded experience. But the ability to deal with it in a detached, non-anxious fashion is also facilitated. This state of mind is similar to a phenomenon reported by patients in psychoanalysis. Associations are experienced as derivatives of one's own mental processes, regardless of the reality of the objects represented. An observing attitude can be maintained until anxiety or other affects become too intense. (Cf. Sterba, 1930.) Sato (1958), too, feels this state of mind may have its counterpart in the free association method of psychoanalysis, but emphasizes the fact that Western therapy usually works with words, whereas the Zen student in this stage does not dissect what he is experiencing with ideational operations. "In Zen the ideas, if they appear, are allowed to pass uncared about. They need not be grasped or verbalized" (p. 217).

Obviously, there are dangers. Traditional Zen literature considered this middle phase to be one in which the monk was in danger of possession by demons (Luk, 1960). There are handbooks, unfortunately not translated, which warn the student against the many experiences in this phase to which he might be tempted to "cling." While the Zen master does not attempt to interpret the emerging material, he is apparently active in guarding the student against acting-out, unconscious projection, and loss of awareness of its subjective origin. He also acts to curb temptations to go into a stupor, toy with paranormal psychic functions, indulge in ecstasies or quietistic retreat. All of these are considered blind alleys. Such dangers must be taken into account in any proposed application of Zen or its procedures. They do not seem to occur during the early portions of training, but later supervision may be critical. Herrigel (1956) does not report having had florid reactions at this stage, and there are apt to be gross individual differences. Where they occur they are considered to be manifestations of the dualistic mind attempting to defend a fictitious ego.

What follows seems to be a very long period of struggle until the intellectual approach is exhausted. Kondo (1952), Herrigel (1956), and Suzuki (1956) describe nearly identical experiences when, having reached a state of inner stillness, the student is given a koan to solve. Time after time he reaches an intel-

lectual solution which is rejected by the Zen master. With increasing despair he concentrates more and more on the koan until his concentration is no longer voluntary; the problem cannot be put aside. Where the hua t'ou is used, the feeling of doubt which emerges is carefully maintained until something like a crisis is reached. This crisis is still rather than turbulent. Intellectual skills seem worthless for the problem, and one can only wait. Death images like "walking in darkness" or "enclosed in a black lacquer casket" are used to describe the state of mind. In the midst of it, the student "lets go" of his egoistic self, "throws himself into the abyss," and satori follows.

INTERPRETATIONS OF SATORI

Erich Fromm, who has conferred extensively with Suzuki, concludes that satori is not a pathological phenomenon. He interprets Zen statements about the limitations of ordinary consciousness in terms of his own explanations of estrangement and alienation. Social learning and individual conflicts produce a "filter" through which only a part of one's experience may pass to be represented in consciousness. He (Fromm et al., 1960) understands satori to be the result of breaking through this filter; "the immediate unreflected grasp of reality, without affective contaimination and intellectualization" (p. 133). In a sense it is a repetition of the direct grasp of the child, but on a new level, in the context of the full development of adult reason, objectivity and individuality. If we define the unconscious to include *all* the aspects of experience which are filtered out, and if we carry the psychoanalytic goal of "making the unconscious conscious" to its ultimate extreme, then the goal of psychoanalysis approximates the Zen goal of enlightenment. The methods obviously differ. To Fromm, Zen makes a frontal attack on the alienated mode of experience, while psychoanalysis trains consciousness to get hold of the unconscious by directing attention, step by step, to individual distortions and personal fictions in the perception of experience.

Schachtel (1959) appears to be in essential agreement about the nature of satori, as a breaking away from conventionalized and overstructured experience. He has traced the role of this mode of experience in adult amnesia for childhood experience and brings forward the death imagery surrounding the presatori state as evidence of the defensive purposes for which the conventionalized mode is used.

Fingarette (1958) has undertaken a fairly comprehensive mapping of mystic states into psychoanalytic ego psychology. His material is taken mainly from the Eastern literature of mysticism including that of Zen. He feels that interpretations which emphasize pathological regression fail to take into account the "significant marks of insight" associated with the great mystics. Patients who have made progress in psychoanalysis describe their changed experience in terms which are sometimes comparable to mystic paradoxes. The mystics may be describing a singularly integrated mode of ego-syntonic functioning. Selflessness cannot refer to an actual loss of the subject-object distinction as it occurs in hallucinations and paranoid delusions, because the social behavior of mystics is often highly realistic and effective. Rather it refers, he feels, to that normal unself-conscious characteristic of experience which is nonanxious and motivated by neutralized drives functioning within the conflict-free portions of the ego. Fingarette (1958) says:

> The introspected, self-conscious "I" is not in fact a perception of one's total person; it is some particular part affect, idea or action of the person as perceived by the person in a context where the dynamically dominant affect is some form of anxiety. Consciousness of self is not an awareness of some self-identical entity; it is, rather, any consciousness colored by intrapsychic conflict and anxiety (p. 16).

"Freedom from striving," "acceptance," and "desirelessness" seem to refer, not to a flaccid absence of motivation, but to absence of inner conflict. Similarly, "no-mind" refers to the absence of compulsive thoughts about thoughts. "Dwelling in voidness" refers to complete openness to experience, unblocked by preconceived ideas or overly-rigid maintenance of logical forms.

Fingarette suggests that such integrated functioning indicates that the mystic has undertaken a prior phase similar to psychotherapy in which he has achieved a deep personality reorganization by facing his underly-

ing problems and gaining insight into them. The regressive phenomena—trances, hallucinations, and so forth—reported in the literature of mysticism appear to belong to an earlier phase of self-exploration rather than to the mystic state itself.

THE FEELING OF ONENESS

In satori one experiences the universe as a totality of being of which oneself and all other objects are manifestations. This is the aspect of satori which is most inexpressible, yet known so directly as to be irrefutable by logical argument.

William James (1928) considered mystic states to be characterized by the subliminal or subconscious mind's breaking through into consciousness. He noted that union, ineffability, certainty, and passive reception were found in most descriptions. While the experience carries strong authority for the individual himself, other people must evaluate it in terms of the empirical functioning of the person after the experience. Mere contact with the subconscious does not guarantee that the experience is psychologically constructive. Certainly the material which James reviewed ranges between extremes of sickness and health. He did not carry his thinking about the subconscious far enough to try to account for which incursions are apt to be pathological and which not.

Psychoanalytic interpretations of oneness begin with Freud's (1955) comparison of the oceaning feeling with the primal experience of unity of the satisfied infant with the maternal breast. This interpretation, in terms of oral fantasies, is followed by many subsequent psychoanalytic writers.

Bertram Lewin (1950) has explored the oral basis of the feeling of union as it occurs during the early phases of manic or hypomanic episodes and in the writings of two Christian mystics. Psychoanalytic material suggests that the earliest feeding experiences of the infant involve not only the wish to devour, but also the wishes to be devoured and to sleep. The well-fed and predormescent infant apparently feels itself merged with and devoured by the breast, and later sleep comes to have this meaning. This is the matrix of subjective experience from which spring ideas of sleep, death, nirvana, immortality, heaven, and the oceanic feeling. The continuing oral meaning of sleep among normal adults may be inferred from hypnogogic phenomena. Typically the sleeper may feel something being pushed into his mouth (devouring) and that something is enveloping him (being devoured).

In the hypomanic episodes of Lewin's patients the wishes to be devoured and to sleep broke through as deeply regressive phenomena leading at times to ecstatic experiences of union. Lewin feels that in these ecstasies the breast is often condensed psychologically with the superego, a deathless one with which the ego identifies so as to participate in its immortality. Along with the active, devouring fantasies there is the sense of yielding, and ultimately joining it in sleep or a sleep-like state. The hypomanic episode may follow and serve to deny and ward off these wishes.

Lewin (1950) makes an interesting contribution by treating the feelings of inexpressibility and certainty as elements of manifest content. He points out that the mystic and his patients have sometimes given excellent descriptions of their experiences. The indescribability is subjective and points to the nonverbal latent content, the union at the breast. Similarly, the noetic quality reflects the realness of the breast experience:

> This experience is what one knows because it is primal, immediate, and unquestioned experience. It was not learned by hearsay or seeing, but represents the primitive narcissistic trust in subjective experience (pp. 149–150).

Fingarette (1958) concedes that the feeling of oneness may be considered a fantasy of the primal unity, but occurring in the context of a highly integrated and flexible ego, and made possible by this flexibility. Acceptable residues of infantile fantasy and partial instinct gratification can be maintained without making experience antirealistic. Thus, the selflessness of anxiety-free experience would, through regression in the service of the ego, be deepened and colored by the selflessness of the primal fantasy. The sense of joy and power derived from conflict-free functioning would have ecstatic overtones of fantasies of primal gratification and omnipotence.

The feeling of oneness and its accompanying characteristics may also occur quite unexpectedly to nonreligious and apparently quite healthy people. James cited several recorded instances.

James also noted that the experience of oneness sometimes occurred under chloroform. More recently Watts (1960) and Van Dusen (1961) have compared experiences under LSD with satori. Both are widely read in Zen, but neither has experienced satori.

The inspirational phase of creative work may also be accompanied by similar experience for some artists. Ehrenzweig (1957) treats this in a highly stimulating discussion of a book by Marianne Milner. The problem of the artist, he says, is to maintain an ego organization flexible enough to gain periodic contact with the less differentiated images outside ordinary consciousness. These images rescue art from stultification by the clichés of surface consciousness. In the "creative surrender" the subject-object distinction may be temporarily abandoned in an oceanic fusion between inner and outer world. Since the ordinary feeling of self is abandoned in this state of consciousness, this surrender may be experienced in sado-masochistic terms as death and rebirth. The affective tone of the experience tends to vary from catastrophic fear to a blissful, almost austere stillness depending on the flexibility and tolerance of the ego.

Ehrenzweig makes an extremely interesting comparison between the ways Schreber and Rilke dealt with undifferentiated imagery stemming from regressed states. Both men apparently experienced bisexual fantasies of procreation in states of consciousness devoid of distinctions between sexes and bodily orifices. Schreber, in order to fend off the ego-disruptive effect of such experience, imposed the precise verbal structuring of surface consciousness on his fantasy. The result was an obscene delusion. Rilke submitted to the undifferentiated state more voluntarily. He was not compelled to articulate his fantasy fully in order to stave off disorganization, but kept its undifferentiated ambiguity in his poetry.

This line of reasoning may help to explain differences in the experience of union as it is reported by different mystics. The "surface" consciousness is faced with the problem of rendering the experience into its own terms. In ordinary consciousness the undifferentiated experience undergoes a kind of "structural repression," the extent of which varies with the tolerance of the ego. A person raised in a culture which recognizes a personal deity is likely to turn to such terms to explain his experience. Individuals may also differ in their need to impose structure, so that, even in cultures where "God" is available, some people will not explain their feeling of oneness in these terms. Finally, the feeling of indescribability may be a residual awareness of the artificial structuring which has been imposed.

We have come now to a very complex situation. The experience of oneness, with its attendant feelings of ineffability, certainty, and passive reception, may occur in many different situations—under the onslaught of deeply regressive oral wishes and of certain drugs, in quiet, reflective states of healthy people, and as a fairly voluntary act of some artists, Zen students and other mystics. Union may be experienced in terms of a highly personalized deity or of an impersonal universe. In affective tone the experience may vary from extreme ecstasy to austere stillness. It should be noted that the satori of Zen is clearly of the impersonal, nonecstatic variety. Ecstasies, like trances, hallucinations, and similar phenomena, are considered blind alleys, simply manifestations of the egoistic mind defending itself.

Two main lines of interpretation have been advanced. One is the classical psychoanalytic explanation in terms of a fantasy of the primal union of the infant with the maternal breast. One line of evidence supporting this view is the oral content given by psychoanalytic patients who have had this experience and the oral terms in which certain Christian mystics have expressed themselves. This genetic explanation refers the feelings of indescribability and certainty to the preverbal nature of the nursing situation which is being "relived." The other explanation is a structural one. The radical change in the state of consciousness is considered to represent, in psychoanalytic terms, a structural one. The radical change in the state of consciousness is considered to represent, in psychoanalytic terms, a structural regression of

the ego. This is more in line with Zen thinking, which construes satori as a breaking through of one state of ("conceptualizing") consciousness to another.

REGRESSION IN THE SERVICE OF THE EGO

Schafer (1958) has organized an excellent review of the psychoanalytic concept of regression in the service of the ego. Satori seems to fit into this class of psychologically adaptive regressions for several reasons. First, discussions of satori repeatedly emphasize flexibility in the use of ego functions. One thinks when the situation requires it, and the intellectual mode may be abandoned when it is unnecessary. Secondly, other types of regressions which might be less adaptive—hallucinations, trances, ecstasies—are consistently rejected as spurious by the Zen master. Thirdly, the states of mind which become possible as a result of satori are clearly used for adaptive purposes. Zen was closely involved in the training of the samurai class in traditional Japanese society (cf. Reischauer & Fairbank, 1958, pp. 547–549): more effective action, rather than monastic withdrawal, was the goal. Finally, the implied increase in energy, and decrease in conflict, inhibiting self-consciousness, and anxiety suggest that satori promotes adaptation.

Since the concept seems relevant, we may use it to increase our understanding of Zen and its training procedures. Schafer has also reviewed some general factors which tend to hamper or facilitate the ability to regress for adaptive purposes. Conditions that interfere seem to center around the unconscious significances or the regressive process. Such meanings as passivity and femininity, sinful and defiant transgression, or magically potent destructiveness have been cited.

The ability to regress is fostered by a secure sense of self able to tolerate momentary blurring of boundaries, a well-developed set of affect signals to guard against getting too close to unassimilable contents, relative mastery of early traumata, moderateness of superego pressures, adequate trust and mutuality in relationships, personal and affective communication to other people, and self-awareness.

We might expect the hampering conditions to become issues in the period of Zen training which precedes satori. It is noteworthy that the Zen literature itself focuses mainly on issues involving the sense of self and the techniques used to defend it.

The sequence of Zen training might easily be conceptualized as a series of "regressed" states, each of which develops functions on which succeeding states must depend. For example, the first phase eventually deepens to a kind of relaxed drowsiness in which primary-process derivitives appear. The ability to deal with them in an accepting fashion, enables the student to get through to the next phase, in which much more "regressed" elements are apparent. Feelings of omnipotence appear to color this stage of "mirror-like" detachment. The nonstriving quality of this state is apparently its chief safeguard against maladaptive reactions, even though further impulse derivatives are probably emerging. This nonstriving is probably a lesson of the previous phase. Satori, in turn, grows out of the second phase and is probably based on certain safeguards which were developed there.

LIMITATIONS OF ORDINARY CONSCIOUSNESS

As we have seen, the Zen literature contrasts the limiting quality of ordinary consciousness with the broader "intuitive" functioning which becomes available as a result of satori. When the swordsman is able to lay aside his ordinary consciousness of himself and his situation and to rely on his "trained unconscious" (i.e., another type of consciousness which is not aware of itself) then he becomes capable of remarkable feats. His perceptual sensitivity becomes broader and his motor response becomes more accurate, less hampered by mediating thoughts.

Research on perceptual sensitivity in states of altered consciousness is clearly relevant. The oldest and best known phenomenon in this area is the increase in accuracy which occurs when a forced choice technique is used (see Eriksen, 1958) in a sensory discrimination task. This essentially means that, when a subject is forced to make a definite comparison, he may correctly discriminate two stimuli above the level of chance even when he feels he is completely guessing. This "guessing state

of mind" may be considered an altered state of consciousness. Eriksen reports that his group has found very marked individual differences in the intensity range between where a subject shows better than chance guessing and the point where he begins to report confidence in his judgments.

Another group of investigators has studied the influence of subliminal stimulation on the products of other states of consciousness: dreams, free associations, and spontaneous imagery. The classical study in this area was conducted in 1917 by Pötzl (see Pötzl, Allers, & Teler, 1960) who exposed pictures tachistoscopically for 1/100 of a second. He had his subjects report verbally and by drawings what they had consciously perceived. He demonstrated that parts of the picture which had not been drawn or described appeared in the manifest content of dreams which the subjects reported next morning. Fisher (1954, 1957) and Shevrin and Luborsky (1958), using much the same procedure as Pötzl, have confirmed this phenomenon. Later studies have demonstrated that "preconsciously registered" parts of subliminally presented pictures may subsequently appear in conscious imagery, free association and hallucinations as well as in dreams. Emergence of these elements seems also to be facilitated by LSD (Fisher, 1956, 1957; Friedman & Fisher, 1960). The following conclusions appear to be supported by these studies: (*a*) an enormous amount of intricate visual material is registered in extremely brief time intervals, such as 1/100 or 1/200 of a second; (*b*) while inaccessible to consciousness in the ordinary sense, the registrations may be demonstrated in their effects on the products of other states of consciousness such as dreams, spontaneous imagery, and free association; (*c*) while the previously unreported elements may be reproduced with photographic accuracy, the dreams and images more often show that numerous transformations and distortions have occurred. These distortions are closely related to characteristics of dream mechanisms or primary process patterns of thinking.

Klein (1959) has explored the implications of these studies in a paper on consciousness. He feels the data warrant a distinction between "registration" and "conscious perception." Perception appears to an experience which is not singular and unvarying from one state of consciousness to another. The problem solving state of consciousness called forth by most laboratory situations tends to limit the range of registrations which may appear in conscious perception. Other states of consciousness, such as those explored in these studies, may draw on other registrations, differently structured.

This interpretation seems directly consistent with the Zen point of view regarding ordinary consciousness. The more limited nature of problem solving consciousness has been explained by ego psychologists in terms of controlling structures which impose distinctions between wish and reality, certain and uncertain, dream and waking perception, and so on. Although the major emphasis has been placed on the role played by these structures in making veridical perception and appropriate response possible, it is conceivable that such structures might unnecessarily limit the range of stimuli available to the person in forming an appropriate response. This is acknowledged by theories concerning regression for purposes of artistic creativity (which may require bypassing the well-trodden schemata of problem solving consciousness), but the point has been little explored with reference to motor responses in complex, even dangerous situations.

A serious problem remains in this formulation. If preconscious registrations of stimuli are subject to such marked distortions, how can they facilitate accurate response? The transformations noted by Fisher are corroborated in studies of incidental stimulation by Pine (1960) and Goldstein and Bartol (1960), where the effects were assessed in the production of stories. Fisher has left open the question whether distortions occur in the process of registration or in the memory traces as they are utilized in the formation of dreams, images, and hallucinations. If the distortions occur only in the process of forming imaginative products, then we need only argue that the swordman's state of consciousness, while certainly "altered," is not necessarily identical with the states whose products have been examined in these studies. The fact that unreported elements were reproduced occasionally with great accuracy in the dreams

of Fisher's subjects suggests that preconscious perception need not be distorted.

The limitation imposed by ordinary consciousness on the flexibility and adaptiveness of motor responses has received less attention. It is clear that there are certain motor habits, such as driving a car or using tools, which operate quite successfully outside the range of focused attention once they have been learned (cf. Hartmann, 1951; Kris, 1952). Consciousness does not necessarily improve function: rather, preconscious responses seem to have definite advantages. It should be noted that a part of the training in many of the Zen "arts" consists in overlearning specific motoric acts so that they may be utilized without thought in response to the intuitive perception of the situation. Of course, the ability to suspend conscious reflection is developed by other means. The acts performed in that state of consciousness are thus quite comparable to this class of automatized functions.

Some of the motor phenomena of hypnosis may be relevant. Weitzenhoffer (1953) concludes that most of these feats may be performed in the waking state with sufficient training, and that endurance one some tasks is often due to suggestions against fatigue. We are warned, then, against expecting miraculous increments in motor capacity under hypnosis. But the occurrence of any increment in performance or decrease in fatigue suggests that thoughts about one's capacity which exist in ordinary states of consciousness may unnecessarily limit what one is able to do. The "second wind" phenomenon may also be relevant as a state of altered consciousness in which enhanced motor performance occurs.

There may exist another class of behavior of which the following is an example:

> A student was working in a factory where extremely hot material was being poured into small portable molds. The temperatures involved were such that, if the molds touched water, a serious explosion would occur. He was sitting in a chair when he saw a mold begin to drop into a puddle. He was at the top of a 12-foot wall when the explosion occurred, but remembers nothing in between. He inferred that he climbed the wall with remarkable speed but had no reflexive awareness of the act.

Unless such reports are simply the product of some distortion, there appears to be a class of emergency behavior involving spectacular motor feats performed without awareness. It has received little attention in the psychological literature, and its very emergency character may have precluded systematic observation.

In summary, the material we have reviewed in this section seems consistent with some of the Zen assertions about the limitations of ordinary waking consciousness in forming adaptive reactions to complex situations. Certain states of consciousnes, such as those that obtain during dreaming or other imaginal productions, seem to make use of stimulus registrations that are not available to the perception of ordinary consciousness, for example, that of a subject in a laboratory situation. Even in a laboratory, nonverbal responses may indicate accurate discriminations in the absence of correct verbal reports, and even verbal reports may show finer discrimination when the subject is forced to choose in spite of strong subjective uncertainty.

On the response side we have less information. Overlearned motor skills may operate best when not specifically attended to, but there is little research concerning the effects of reflective consciousness on motor efficiency. Some of the motor phenomena under hypnosis suggest that ordinary consciousness may unnecessarily limit response because of the way one conceptualizes his capacity. Finally, we have a mere hint that much more dramatic motor activities may take place, without conscious reflection, in extreme emergency situations.

TRAINING PROCEDURES

We turn now to consider the basic exercises used in Zen training. Hitherto we have noted some similarities between satori and other, better known states of consciousness. These included mystical and creative "regressions" occurring in the context of well-integrated personalities. While in Western observers these transient episodes seem to have occurred unexpectedly, almost by accident, Zen uses methodical training procedures for bringing about satori. The basic exercise used for this purpose consists of motionless sitting and some sort of concentration. For ease of exposition we will refer to this practice by its Japanese name, *zazen*.

Zazen and Relaxation

Sato (1958) has compared the bodily adjustment of zazen to Jacobson's "differential relaxation," and it is useful to survey the literature related to relaxation and its therapeutic effects. First let us review just what is done with the body during zazen. The ordinary position used in China and Japan is the cross-legged "full lotus" posture seen in many Buddhist statues. The back is straight, the head erect, the hands folded in the lap, and the eyelids partly open. The position is balanced, but not completely relaxed. Most Westerners cannot cross their legs this way without strain, but some Zen authorities feel that the same kind of balance can be achieved sitting on a chair. (Kondo, 1958; Senzaki & McCandless, 1953)

Present-day techniques of therapeutic relaxation training derive mainly from two sources: Jacobson's "progressive relaxation" (1938) and Schultz's "autogenic training" (Schultz & Luthe, 1959). There are other methods, but these two are the most widely used and represent most of the variation to be found in this field. Jacobson's procedures focus on training the subject to recognize muscular tension, this being considered a necessary step toward enabling him to relax it. One method is to have him contract muscle groups in successively smaller degrees until he is aware of even small amounts of tension. With this greater awareness of tension, the subject is told to relax to a greater and greater degree. Toward the end of a training session the subject may fall asleep. Jacobson considers this a sign of successful progress. Further supervision aims at counteracting habits of muscular tension of which the trainee is not aware. After the subject has learned to achieve deep relaxation in a reclining position, he is taught "differential relaxation" in other positions.

While Jacobson's method is a direct approach to relaxation through training in the muscle sense, some other methods are more indirect and require no such conscious control of the musculature. Schultz's autogenic training, which derived originally from autohypnotic techniques, is one of these. This method, while well known and widely used in Germany, is virtually unknown in this country. The first presentation of the method in English (Schultz & Luthe) appeared only in 1959, but it lists over 600 clinical and research reports on its use during the past thirty years.

The method is as follows: the subject sits or lies down and assumes a relaxed state of mind. He concentrates on a "training formula," such as "my right arm is heavy" while maintaining mental contact with that portion of his body. An attitude of "passive concentration"—of focused awareness but unconcerned indifference to the functional outcome—is important in producing effects. Initial sessions may be as short as 5 minutes with 1-minute periods of concentration punctuated by 1-minute periods of rest. The full standard series of formulae proceeds from heaviness in one arm, then both arms, then both legs, to warmth in the extremities, "heartbeat calm and regular," "it breathes me," "my solar plexus is warm," and "my forehead is cool." Each formula is added only after preceding ones are well established.

Since the kind of mental contact involved has very real consequences for the circulation, blood pressure, and so on in the body parts concerned, careful supervision is maintained to avoid undesirable physical reactions. These are primarily of two kinds: (*a*) the trainee's mental contact may be incorrect (for example, the hand rather than the whole arm may be the area of focus) in which case a variety of minor symptoms may appear; (*b*) in the case of special irritability of certain body parts, as in psychosomatic conditions, serious complications may occur unless formulae involving these areas are approached carefully. Thorough mastery of the series ordinarily takes from one to eight months. Schultz and Luthe (1959) say:

> The trainee should be capable of "switching on" the many various standard exercises effectively and almost instantly by applying the following pattern of formulae: "I am at peace . . . My arms and legs are heavy . . . I am at peace . . . Heartbeat calm and regular . . . It breathes me . . . My solar plexus is warm . . . My forehead is cool." As a result of the quick (20–30 seconds) and effective application of the standard formulae, most trainees will experience the body as a resting mass which is heavy and warm. Slow pulsation of the heart and deep and slow respiration may be perceived. The head is usually experienced as being "separated" from the rest of the body (pp. 95–96).

What are the effects of relaxation? In a series of studies using relaxation trained subjects Jacobson (1930a, 1930b, 1930c, 1930d, 1931a, 1931b) presented evidence that mental activity is impossible without minute quantities of muscular tension. Taking pains not to suggest his main interest in these studies, he had subjects visualize various objects. Regularly they reported minute quantities of tension in the muscles of the eye. Told to relax completely while visualizing objects, the subjects could do one or the other, but not both at the same time. Using a string galvanometer to measure muscle potentials, he demonstrated similar effects with imagined movements. Since this evidence indicates that mental and emotional activity are associated with neuromuscular activity, Jacobson (1941) conceived of relaxation training as providing a direct, mechanical, negative control over such functions and hence as a sufficient treatment in many neuroses. While this formulation leaves important questions unanswered, the evidence does suggest that relaxation may affect the quantity of mental activity and that some measure of control may be exerted by relaxation procedures.

Another effect of relaxation procedures is less well documented and more subjective, but it is worth examining. This is the possibility that by focusing attention on the present state of the body, relaxation techniques may indirectly act to lessen anxiety. Thinking tends to take the subject away from his immediate feelings and sensations: one can only anticipate the future, remember the past, or deal with distant objects cognitively. It is generally agreed that anxiety is an anticipation, even though the content (such as bodily destruction or loss of objects) may not be conscious in neurotic anxiety. Being anticipatory, it is cognitive. All of the relaxation methods reviewed seek in some way to heighten awareness of the present state of the body. This awareness, in a sense, brings the subject back to a personal "here and now," away from the the thought-world of distant objects and events. Thus it is reasonable to expect that cognitive elaborations of anxiety may be cut short. The means by which relaxation training lessens pain may, in part, be similar. Full relaxation may not be necessary to achieve these results. Jacobson, later in his life (1955), reported decrease in anxiety and greater feelings of control when subjects were trained to attend to the sensations associated with acts rather than attending simply to the goals sought.

A third immediate effect of relaxation may be cathartic release of thoughts or emotions of which the subject was previously unaware. This effect is not invariable, but seems to occur in a certain proportion of cases. It may be partly due to psychic content related to the tension itself. Jacobson (1938) was inclined to believe that muscular tension always had to do with some act—so much so that he was inclined to ask his trainees what they felt themselves tensing to *do*. Defensive conflicts may easily find expression in muscular tension (cf. Fenichel, 1945, p. 246). Braatøy (1942; 1952; 1954, pp. 155–197) has discussed habitual patterns of muscular tension in neurotics as more generalized techniques of suppressing emotions. Relaxation on the psychoanalytic couch releases spontaneity.

The point is that muscular tensions often, if not always, express something. In certain individuals, the pattern of muscular tension may be involved in habitual, automatic, constant attitudes maintained to suppress emotion or to fend off the motor expression of some impulse. These attitudes will interfere with the achievement of relaxation. When relaxation is achieved anxiety, emotional catharsis, or emergence of previously repressed ideas may occur. The relationship between muscular tension and psychic content has not been studied sufficiently to enable us to predict which individuals will react in this way. Hadley (1938) reports that catharsis in a physically relaxed patient is more spontaneous and is not accompanied by the bursts of emotion which characterize ordinary catharsis.

Relaxation procedures appear to differ in the psychic states which they induce. Jacobson's techniques seem to lead to a "blank mental state" and frequently to sleep. It is noteworthy that the Zen literature warns the student against this blank state during zazen (Hui Hai, 1948; Luk, 1960). It is felt to be useless and to hinder the occurrence of satori. Bagchi (1936) notes that Hindu theories of relaxation aim at quiet attention, and later to a state of attenuated conscious-

ness which they consider a different state of equilibrium from sleep. He suggests that Jacobson makes use of attitudes similar to quiet attention, but his mechanistic outlook causes him to overlook their importance in inducing relaxation. Autogenic sessions do not necessarily lead to sleep.

The long-range effects reported by trainees are remarkably uniform regardless of the method used. The pattern includes quicker and more restful sleep, feelings of increased energy and endurance, less tension and greater ability to cope with difficult situations, less susceptibility to emotional outbursts, and a greater sense of self control. This last effect—the reassurance that one has an active means of controlling tension and difficult feelings—may be especially important.

Observations of behavioral changes attributable to relaxation primarily concern reduction in symptoms of anxiety and tension. Motor effects were studied by Haverland (1953), who gave biweekly training in progressive relaxation for 6 weeks to 26 subjects. Control groups received training in rhythmic exercises or ordinary physical education. The relaxation group showed significantly greater improvement on a reaction-time task, and on several tasks related to motor control (steadiness, aiming, tracing). Improvement in motor control and decreased muscular tension may also be involved in the results reported by Neufeld (1951). A group of several hundred Naval Air Cadets were given group training in progressive relaxation for 5 weeks—a very short course. This group had significantly fewer physical injuries as well as fewer days lost for other sickness than similar groups of nontrained cadets. A smaller group was systematically observed for several nights while sleeping. Compared to a nontrained control group these subjects tended to fall asleep faster, move about and waken less often, although these differences were not significant. Yates (1946) reported good results using relaxation training to improve performance of college boxers and Army aviators, although World War II intervened before sufficient follow-up study could be carried out.

Where relaxation training has been used in conjunction with psychotherapy, therapists report that patients associate more freely (Hadley, 1938; Pascal, 1947; Schade, Hruza, Washburne, & Carns, 1952; Schultz & Luthe, 1959). Autogenic training in particular has frequently been used in conjunction with psychoanalysis. One interesting effect is that patients subject to anxiety dreams report these dreams begin to take on benign endings. Most investigators have found relaxation training alleviates symptoms of tension and promotes insight into psychosomatic relationships. This symptomatic relief may leave patients better able to attack underlying problems in therapy. But several observers warn that the symptomatic relief may decrease motivation to continue treatment in some cases.

Much more extensive observation has been made on the effects of relaxation on psychosomatic disorders. This is especially true of autogenic training. The range of disorders treated and the varying effectiveness of the treatment in different disorders would be impossible to cover in detail here. Autogenic training is often used in the treatment of bronchial asthma. A tendency for autogenic exercises to correct deviations in blood sugar balance has been demonstrated in well-controlled studies. Diabetics often require marked reductions in insulin dosage. A wide range of milder psychosomatic disorders, such as cardiac neurosis, neurotic epigastric symptoms, and gastritis, respond well. In some other disorders, where the effectiveness of autogenic training has not been established, there have been isolated instances of striking improvement. Some epileptic patients report being able to stave off seizures as a result of training, although the frequency of aurae remains the same. Jacobson (1920, 1938, 1940) reported cases of hypertension, mucous colitis, and other disorders which responded well to progressive relaxation. His case reports, while covering very few patients, include excellent follow-up studies—some for 10 years or more.

Most authorities agree that the applicability of relaxation training depends largely on the capacity of the individual to take responsibility for applying the exercises. Children below about 9 and mentally retarded individuals do not seem to respond well. Neither do people who lack self-direction, either by illness or excessive dependent needs. Beyond this limitation, the type of disorder involved need not preclude benefits. Schade et al. (1952) were unable to find clear-cut

diagnostic groups that did not respond. There are, however, some suggested differences in response between different types of neurotics. Hysterics respond quickly at the outset of autogenic training, but tend to have trouble practicing without supervision. Obsessive-compulsive patients take a long time to establish the initial formulae in autogenic training, but past the initial stage they respond well (Schultz & Luthe, 1959).

It seems clear from this review that from a physical standpoint, the meditative sitting of Zen may be subsumed under the category of relaxation training. Nearly all of the methods extend relaxation training to some sitting position, and the majority of them use similar breathing exercises to facilitate relaxation. From this literature, then, we may gain information about the probable effects of zazen in its early stages. This includes the pattern of subjectively felt benefits, possible physiological reactions, and so on. But in all of this literature, there is no mention of any experience like satori. Zen Buddhists evidently regard meditative sitting as a means of suspending ordinary conceptual activity. Jacobson's evidence suggests that relaxation may indeed alter mental activity. The difference between zazen and these other relaxation techniques probably lies in the use to which the relaxed state of mind is put.

Concentration and Zazen

The other aspect of zazen, and the one which receives the most comment in Zen literature, is an attitude of concentration. This is not a strained or effortful focusing of attention, but a state of mind quite similar to what Schultz calls "passive concentration"—a peaceful attention to the object of concentration without straining to achieve effects. Relaxation itself, by decreasing mental activity, appears to facilitate this state of mind. Extraneous distractions—thoughts and external stimuli—are not forcibly excluded but rather dismissed. Eventually a state of stillness ensues which is occasionally broken by spontaneous associations and feelings.

There appears to be good introspective evidence that focused attention as it is ordinarily deployed may serve to limit the form and quality of conscious contents. The reorderings and transformations necessary at certain stages of creative thinking (and in therapeutic insight) seem to occur most readily at the "fringes" of focal awareness. The "freely floating attention" which enables the psychoanalyst to "hear" his patient on many different levels seems to involve a change in ordinary attention. Titchener (1912) noted that when attention is directed at affects they seem to evaporate, to lose their reality. When attention remains on the object of the affect, the affect remains vivid. There thus seems to be a close relationship between the logical, nonemotional schemata of adult waking consciousness and attention deployment.

Since attention deployment seems to be related to the logical schemata of adult waking consciousness, it is reasonable to expect that when attention is focused on one "corner" of the conscious field differently organized content will emerge. Recent studies of perceptual isolation, where attention deployment is altered artifically by restricting afferent stimulation, seem to support this (Solomon, 1961). Imagery, primary process thought, and altered bodily feelings tend to occur quickly, although with marked individual differences in tolerance and in the patterns of alteration. It is tempting to explore the perceptual isolation literature further in search of possible insights into zazen. However, the differences seem more important than the similarities. The Zen student sets out to deal with mental activity in a fairly specific way, whereas the perceptual isolation subject is left pretty much to his own devices. The relaxation aspect of zazen is likely also to introduce important differences in the amount of anxiety, press of ideation, and affect which occur in the two situations.

As the spontaneous associations emerge in zazen, the student strives to maintain a detached view of them, without acting out or otherwise distorting reality in terms of them. He simply accepts and observes them until they pass. The Zen literature does not tell us specifically what the Zen master does to aid the student to maintain his detachment. Possibly the Zen master may not need to deal with "resistances" as actively as the psychoanalyst, because of the nature of zazen. If so, one factor may be that concentration is a less complex task than free association:

the student may more easily become aware when he is deviating from the task. To put it another way, the attempt not to think may be particularly effective in helping the subject to be aware of the subjective origin of his thoughts and feelings. Another factor may be the increased awareness of the personal present which was discussed earlier.

There is some research indicating individual differences in attention deployment which could influence how different persons will respond to zazen. This is the so-called "scanning control" principle which has emerged in studies by Schlesinger (1954) and Gardner, Holzman, Klein, Linton, and Spence (1959). These studies suggest that certain individuals tend to scan the stimulus field, or deploy attention broadly, while others tend to focus their attention more narrowly. The patterns seem to appear consistently, as characteristics of the individual in a variety of situations. The research also suggests a link between broad attentional scanning and the defense mechanism of isolation, although broad scanning does not seem to be used for defensive purposes by all the subjects observed. Inasmuch as Zen training requires a more narrow focusing of attention, individuals who scan broadly might have special difficulty in carrying out the procedures. Where broad scanning is involved in defensive patterns it should be particularly difficult to suspend. The relevance of this cognitive control principle to performance in situations similar to Zen meditation remains to be explored.

THERAPEUTIC APPLICATIONS

In order to explore possible therapeutic applications of Zen Buddhism we must distinguish between the experience of satori and Zen meditation. Satori seems to facilitate very healthy personality functioning. But satori takes years of training and probably a Zen master as well. For practical reasons its applicability to Western psychotherapy seems limited. Most present-day applications of Zen to psychotherapy derive from zazen.

Kondo (1958) reports that he instructs his neurotic patients to practice sitting and breath concentration in addition to their psychoanalytic sessions with him. After an initial exacerbation of symptoms the exercise seems to aid progress in therapy. Activity which has served as an escape mechanism to avoid facing problems is blocked by the sitting. A more unified feeling of self and a calm vigor is the eventual result. Gradually the patient "begins to show, unconsciously, more intensive concentration in working on his problems in the therapeutic situation. . . . His psychic energy has begun to become assembled, unified, and available for constructive work." Kondo does not report individual differences in the reactions to this application. Fromm (1959) has suggested that some such application might be particularly helpful in the treatment of character disorders.

There have been several widespread applications of zazen in modern Japan. According to one very indirect source (Brower, 1961) Tokyo bus drivers have been required to practice zazen in recent years. The reported result is a decrease in street accidents involving busses. Many Japanese military officers are said to have received training in zazen before and during World War II (Dewey, 1920; Malm, 1959). Unfortunately, I have not yet found more direct information about these applications and their results.

Morita therapy, developed by a Japanese physician of that name, has recently been cited as a psychotherapy with a viewpoint akin to that of Zen (Kora & Sato, 1958). The treatment begins with several days of complete bed rest without distraction. The patient is simply instructed to leave himself to the "dynamics of the situation": to eat and sleep as he wishes, and to accept feelings as they come. As the patient permits himself to suffer, worry, be uncomfortable without resisting, his discomfort gradually recedes. After 4 or 5 days a feeling of ennui appears, and the patient is starved for stimulation. He is given simple chores. Gradually his work becomes more complex until he is able to return to his ordinary job. The whole treatment usually takes from 4 to 5 weeks. The treatment is felt to require considerable ego strength, and it is used mainly with neurasthenic and obsessional patients. Results at the Kyushi University Medical School are a reported 76% cured, 8% improved. No follow-up information is reported.

Humphreys (1960) reports that his group of English subjects who practiced sitting and breath concentration for several years "in-

creased in intuitive development," withdrew projections by which they had tended to distort reality, showed greater serenity, ability to cope, and compassion. "All who have made this experiment in the last few years have changed remarkably, passing of course through periods of depression and doubt, but finding these well suffered as the price of wider awareness, deeper understanding of eternal truths, and many a brief experience of things no words can usefully describe" (p. 205).

In closing we must raise the possibility that the use of zazen may enable patients to deal with problems which are inaccessible to other kinds of psychotherapy. A number of psychoanalytic writers have pointed to such a class of problems. Balint's (1958) discussion is particularly broad in scope. The usual argument is that problems stemming from Oedipal sources are most amenable to psychotherapy. The Oedipal experiences, regressively resurrected in the transference, are susceptible to verbal report. Verbal interpretations are understood as such and can be worked through. Defenses at this level enable the patient to internalize despite tension. But in the context of problems related to pre-Oedipal experience, the verbal communication of therapy runs into more difficulty. Words do not always have an agreed, conventional meaning. Externalizing defenses such as acting out and projection correspond to pre-Oedipal levels and are less amenable to therapeutic handling. In the self-exploration of zazen there is no need to verbalize the emerging experiences, yet anxiety and the need to externalize seem to be decreased. Might it not be that subjects could learn to face and accept experience corresponding to these lower levels by using zazen? Satori itself, with its marks of preverbal, prelogical experience, seems to affirm this possibility.

REFERENCES

BAGCHI, B. K. Mental hygiene and the Hindu doctrine of relaxation. *Ment. Hyg., N. Y.*, 1936, **20**, 424–440.

BALINT, M. The three areas of mind. *Int. J. Psychoanal.*, 1958, **39**, 328–340.

BRAATØY, T. The neuromuscular hypertension and the understanding of nervous conditions. *J. nerv. ment. Dis.*, 1942, **95**, 550–567.

BRAATØY, T. Psychology versus anatomy in the treatment of arm neuroses with physiotherapy. *J. nerv. Dis.*, 1952, **115**, 215–245.

BRAATØY, T. *Fundamentals of psychoanalytic technique*. New York: Wiley, 1954.

BROWER, MILLICENT. Article on the First Zen Institute. *The Village Voice*, January 5, 1961.

CHANG, CHEN-CHI. *The practice of Zen*. New York: Harper, 1959.

DEWEY, J. Letter of April 1, 1919. In, *Letters from China and Japan*. New York: Dutton, 1920.

EHRENZWEIG, A. The creative surrender. *Amer. Imago*, 1957, **14**, 193–210.

ERIKSEN, C. W. Unconscious processes. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1958*. Lincoln: Univer. Nebraska Press, 1958. Pp. 169–227.

FENICHEL, O. *The psychoanalytic theory of neurosis*. New York: Norton, 1945.

FINGARETTE, H. The ego and mystic selflessness. *Psychoanal. psychoanal. Rev.*, 1958, **45**, 5–40.

FISHER, C. Dreams and perception: The role of preconscious and primary modes of perception in dream formation. *J. Amer. Psychoanal. Ass.*, 1954, **3**, 389–445.

FISHER, C. Dreams, images and perception: A study of unconscious-preconscious relationships. *J. Amer. Psychoanal. Ass.*, 1956, **4**, 5–48.

FISHER, C. A study of the preliminary stages of the construction of dreams. *J. Amer. Psychoanal. Ass.*, 1957, **5**, 5–60.

FREUD, S. *Civilization and its discontents*. (Originally published 1930) London: Hogarth, 1955.

FRIEDMAN, S. M., & FISHER, C. Further observations on primary modes of perception: The use of a masking technique for subliminal visual stimulation. *J. Amer. Psychoanal. Ass.*, 1960, **8**, 100–129.

FROMM, E. Psychoanalysis and Zen Buddhism. *Psychologia*, 1959, **2**, 79–99.

FROMM, E., SUZUKI, D. T. & DEMARTINO, R. *Zen Buddhism and psychoanalysis*. New York: Harper, 1960.

GARDNER, R. W., HOLZMAN, P. S., KLEIN, G. S., LINTON, HARRIET P. & SPENCE, D. P. Cognitive control: A study of individual consistencies in cognitive behavior. *Psychol. Issues*, 1959, **1**(4), 1–186.

GOLDSTEIN, M., & BARTOL, R. P. Fantasy responses to subliminal stimuli. *J. abnorm. soc. Psychol.*, 1960, **60**, 22–26.

HADLEY, J. M. Various roles of relaxation in psychotherapeutics. *J. gen. Psychol.*, 1938, **19**, 191–203.

HARTMANN, H. Ego psychology and the problem of adaptation. In D. Rapaport (Ed.), *Organization and pathology of thought*. New York: Columbia Univer. Press, 1951.

HAVERLAND, L. E. H. The effects of relaxation training on certain aspects of motor skill. Unpublished doctoral dissertation, University of Illinois, 1953.

HERRIGEL, E. *Zen in the art of archery*. New York: Pantheon, 1953.

HERRIGEL, E. *The method of Zen*. New York: Pantheon, 1960.

HUI HAI. *The path to sudden attainment*. London: Buddhist Society, 1948.

HUMPHREYS, C. *Zen comes west: The present and future of Zen Buddhism in Britain*. New York: Macmillan, 1960.

JACOBSON, E. Reduction of nervous irritability and excitement by progressive relaxation. *J. nerv. ment. Dis.*, 1920, **53**, 282.

JACOBSON, E. Electrical measurements of neuromuscular states during mental activities: I. Imagination of movement involving skeletal muscle. *Amer. J. Physiol.*, 1930, **92**, 567–608. (a)

JACOBSON, E. Electrical measurements of neuromuscular states during mental activities: II. Imagination and recollection of various muscular acts. *Amer. J. Physiol.*, 1930, **94**, 22–34. (b)

JACOBSON, E. Electrical measurements of neuromuscular states during mental activities: III. Visual imagination and recollection of muscular acts. *Amer. J. Physiol.*, 1930, **95**, 694–702. (c)

JACOBSON, E. Electrical measurements of neuromuscular states during mental activities: IV. Evidence of contraction of specific muscles during imagination. *Amer. J. Physiol.*, 1930, **95**, 703–712. (d)

JACOBSON, E. Electrical measurements of neuromuscular states during mental activities: V. Variation of specific muscles contracting during imagination. *Amer. J. Physiol.*, 1931, **96**, 115–121. (a)

JACOBSON, E. Electrical measurements of neuromuscular states during mental activities: VI. A note on mental activities concerning an amputated limb. *Amer. J. Physiol.*, 1931, **96**, 122–125. (b)

JACOBSON, E. *Progressive relaxation.* (2nd ed.) Chicago: Univer. Chicago Press, 1938.

JACOBSON, E. Cultivated relaxation in essential hypertension. *Arch. phys. Ther.*, 1940, **21**, 645–654.

JACOBSON, E. The physiological conception and treatment of certain common "psychoneuroses." *Amer. J. Psychiat.*, 1941, **98**, 219.

JACOBSON, E. Neuromuscular controls in man: Methods of self direction in health and disease. *Amer. J. Psychol.*, 1955, **68**, 549–561.

JAMES, W. *The varieties of religious experience.* (Originally published 1902) New York: Longmans, Green, 1928.

KLEIN, G. S. Consciousness in psychoanalytic theory. *J. Amer. Psychoanal. Ass.*, 1959, **7**, 5–34.

KONDO, A. Intuition in Zen Buddhism. *Amer. J. Psychoanal.*, 1952, **12**, 10–14.

KONDO, A. Zen in psychotherapy: The virtue of sitting. *Chicago Rev.*, 1958, **12**(2), 57–64.

KORA, I., & SATO, K. Morita therapy: A psychotherapy in the way of Zen. *Psychologia*, 1958, **1**, 219–225.

KRIS, E. *Psychoanalytic explorations in art.* New York: International Univer. Press, 1952.

LEWIN, B. D. *The psychoanalysis of elation.* New York: Norton, 1950.

LUK, C. *Ch'an and Zen teaching.* London: Rider, 1960.

MALM, W. P. *Japanese music and musical instruments.* Tokyo: Tuttle, 1959.

NEUFELD, W. Relaxation methods in U. S. Navy air schools. *Amer. J. Psychiat.*, 1951, **108**, 132–137.

PASCAL, G. R. The use of relaxation in short-term psychotherapy. *J. abnorm. soc. Psychol.*, 1947, **53**, 226–242.

PINE, F. Incidental stimulation: A study of preconscious transformations. *J. abnorm. soc. Psychol.*, 1960, **60**, 68–75.

PÖTZL, O., ALLERS, R., & TELER, J. Preconscious stimulation in dreams, associations, and images: Classical studies. *Psychol. Issues*, 1960, **2**(3, Monogr. No. 7).

REISCHAUER, E. O., & FAIRBANK, J. K. *East Asia: The great tradition.* Boston: Houghton Mifflin, 1958.

SARTRE, J. P. The transendence of the ego: An existentialist theory of consciousness, New York: Noonday, 1957.

SATO, K. Psychotherapeutic implications of Zen. *Psychologia*, 1958, **1**, 213–218.

SATO, K. How to get Zen enlightenment: On Master Ighiguru's five days intensive course for its attainment. *Psychologia*, 1959, **2**, 107–118.

SCHACHTEL, E. G. *Metamorphosis.* New York: Basic Books, 1959.

SCHADE, M., HRUZA, T., WASHBURNE, A., & CARNS, M. Relaxation as an adjunct to psychotherapy. *J. clin. Psychol.*, 1952, **8**, 338–346.

SCHAFER, R. Regression in the service of the ego: The relevance of a psychoanalytic concept for personality assessment. In G. Lindzey (Ed.), *Assessment of human motives.* New York: Rinehart, 1958.

SCHLESINGER, H. J. Cognitive attitudes in relation to susceptibility to interference. *J. Pers.*, 1954, **22**, 354–374.

SCHULTZ, J. H., & LUTHE, W. *Autogenic training: A psychophysiologic approach in psychotherapy.* New York: Grune & Stratton, 1959.

SENZAKI, N., & MCCANDLESS, RUTH S. *Buddhism and Zen.* New York: Philosophical Library, 1953.

SHEVRIN, H., & LUBORSKY, L. The measurement of preconscious perception in dreams and images: An investigation of the Poetzl phenomenon. *J. abnorm. soc. Psychol.*, 1958, **56**, 285–294.

SOLOMON, P. (Ed.) *Sensory deprivation: A symposium held at Harvard Medical School.* Cambridge, Mass.: Harvard Univer. Press, 1961.

STERBA, R. The fate of the ego in analytic therapy. *Int. J. Psychonanal.*, 1930, **11**, 12–23.

SUZUKI, D. T. *Living by Zen.* Tokyo: Sanseido, 1949. (a)

SUZUKI, D. T. *The Zen doctrine of no-mind.* London: Rider, 1949. (b)

SUZUKI, D. T. *Zen Buddhism.* Garden City: Doubleday, 1956.

TITCHENER, E. B. *A textbook of psychology.* New York: Macmillan, 1912.

VAN DUSEN, W. LSD and the enlightenment of Zen. *Psychologia*, 1961, **4**, 11–16.

WATTS, A. *This is it, and other essays on Zen and spiritual experience.* New York: Pantheon, 1960.

WEITZENHOFFER, A. *Hypnosis: An objective study in suggestibility.* New York: Wiley, 1953.

YATES, D. H. Relaxation in psychotherapy. *J. gen. Psychol.*, 1946, **34**, 213–38.

(Received November 2, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 379–382

NOTES AND COMMENT

TRANSPARENCY IN THE EDWARDS PERSONAL PREFERENCE SCHEDULE

MAURICE KORMAN AND FRANCES COLTHARP

Southwestern Medical School, University of Texas

Much interest has been generated in recent years in what seem to be undesirable test-taking influences in psychometric instruments such as distortions due to a variety of response sets (Broen & Wirt, 1958; Cronbach, 1950). Too often, sight has been lost of the fact that the crucial question still remains validity, pro and con, as Gustad (1956), Barron (1959), and Hathaway (1960) have recently reminded us. It is nonetheless worthwhile to inquire into possible extratest factors that may account for test variance, especially if a psychometric instrument is constructed largely around the concept of avoiding or minimizing such influences.

The Edwards Personality Preference Schedule (EPPS) is a face-valid test, the construction of which suggests greater concern with the control of undesirable test-taking attitudes, particularly social desirability, than with empirically-derived validity in the traditional sense (Edwards, 1959). Because of this emphasis, the EPPS would become all the more vulnerable to undesirable test taking distortions should these built-in controls prove not to be effective since it is not possible, as in empirically derived tests, to fall back upon the ineluctable fact that a certain set of items—whether loaded with social desirability or not—is a proven validity carrier.

A first issue of interest concerns the transparency of the EPPS items. How readily can a naive subject "guess" what a particular item is designed to measure? A priori, or face-valid tests, tend to be transparent to a larger extent than the more "subtle" portions of empirically derived tests (see Seeman, 1956) since no dramatic superiority in either the theory of verbal behavior or in personology distinguish psychologists from other mortals. If, as we hypothesize, the "meaning" of EPPS items can readily be guessed by individuals for whom it was designed, it would further be expected that the resulting 15-scale profile bears some relationship to the public and socially acceptable personality which the subject feels free to exhibit. Two techniques will be utilized to test this possibility. First, does a relationship exist between the EPPS scales and their social desirability? Navran and Stauffacher, in a much quoted paper published in 1954, suggest that no such relationship exists. Their study is marred, however, by a serious methodological error: EPPS *raw* scores instead of *T* scores were used.[1] Secondly, how does EPPS performance relate to self-report? If the EPPS is relatively transparent, we would expect, as some preliminary results suggest (Mann, 1958), that such a relationship exists.

PROCEDURE

Subjects were 98 freshmen medical students, with no former knowledge of the EPPS. During the early part of their freshman year, the standard EPPS had been given as part of a routine battery of tests. Two months later, each subject was given a list of the 15 personality traits for which the EPPS is scored, with a brief description of each as derived from Murray's (1938) needs, and a second list containing the 135 items which appear in the EPPS questionnaire. Considerable care was taken to avoid any duplication in phrasing be-

[1] As can be seen from Edwards' own norms (1959, p. 14) the raw score distributions of the individual scales are clearly not comparable to one another. The lack of reference to a normative population when making interscale comparisons allows for the influence of random error related to choice, order, wording, etc. of items. On the student norms, for instance, a raw score of 15 yields *T* scores of 48 for Ach, 61 for Def, 52 for Exh, etc. Ranking the scales according to their raw scores would thus introduce considerable error and bear little relationship to parallel *T* score ranks. It should be noted that at the time of the Navran and Stauffacher study, *T* scores for the EPPS were not yet available (Navran, Personal communication, 1960).

tween the need descriptions and the EPPS items. After briefly studying the list of personality traits, the subjects matched the items to the traits on the basis of what seemed to them to be the most appropriate assignments. Additional motivation was provided by offering five-dollar rewards to the two subjects making the fewest errors in matching.

The second task was to have the subjects list the personality traits from 1 to 15 in the order of their social desirability in others. Finally, the subjects were asked to rate themselves on each of the traits, on a nine-point scale.

Four measures were thus available for each subject: (*a*) the standard EPPS profile; (*b*) a transparency score, defined as the number of items correctly matched to the appropriate personality traits; (*c*) a ranking of the traits in the order of their social desirability in others; and (*d*) a self-rating on each trait.

RESULTS AND DISCUSSION

Our subjects obtained a mean transparency score of 104.0, with a standard deviation of 10.5. This means that on the average they were able to correctly match 77% of the 135 items making up the EPPS with the appropriate EPPS scale, a significant departure from chance expectation ($\chi^2 = 68.5, p. < .001$) which clearly demonstrates that there is little subtlety in these items for the average naive student.

It is of interest to look at the four items that were *least* transparent for our subjects: "I like to avoid responsibilities and obligations" (Aut); "I like to forgive my friends who may sometimes hurt me" (Nur); "I like to find out what great men have thought about various problems in which I am interested" (Def); "I like to argue for my point of view when it is attacked by others" (Dom). Paradoxically, it would seem that when we do encounter "subtler" items in a test, the items for which were selected on a face-valid basis, we are forced to question the relevance of the content of these items to the traits they are presumed to measure! From Table 1 it can also be seen that scales such as Order and Dominance are considerably more transparent than scales such as Achievement and Deference.

With regard to the relationship between the subjects' average EPPS scale scores and their average social desirability, we found a considerably greater relationship than did Navran and Stauffacher (1954). Where they report a rho of $-.01$ between the composite EPPS scores and the composite social desirability, we found a rho of .46 ($p < .05$, one-tailed test). As noted above, our data were based on *T* scores whereas theirs were derived from raw scores, an inappropriate procedure in view of the considerable interscale differences in standard score transformations. It would

TABLE 1

TRANSPARENCY, SOCIAL DESIRABILITY, AND SELF-REPORT IN RELATION TO EPPS SCALES

Scale	Rank order of transparency	Mean social desirability rank	Mean self-rating	Mean EPPS *T* scores	Correlations of EPPS and self-ratings
Dominance	2	8.1	4.41	47.00	.04
Abasement	5	12.7	7.00	48.46	.24*
Nurturance	13	7.4	5.04	48.71	.37**
Change	8	11.4	6.40	47.76	.44**
Endurance	9	3.4	3.59	54.90	.26*
Heterosexuality	3	5.6	2.78	55.75	.36**
Aggression	11	12.2	5.85	50.55	.23*
Achievement	14	2.6	2.79	50.08	.16
Deference	15	6.4	4.74	52.15	.20
Order	1	5.6	4.01	51.45	.52**
Exhibition	7	12.0	5.79	49.53	.26*
Affiliation	6	3.6	3.61	47.61	.35**
Autonomy	12	7.2	4.41	48.74	.43**
Intraception	10	9.5	4.40	50.18	.26*
Succorrance	4	12.6	7.25	45.49	.37**

* $p < .05$.
** $p < .01$.

seem therefore that in spite of Edwards' attempts to control for social desirability at the *item* level, there is a significant relationship at the variable level between the relative profile position of a scale and its relative social desirability value.

A parallel situation exists when we examine the relationship between the EPPS and self-report. We found a rho of .56 ($p < .05$, one-tailed test) between the composite EPPS score ranking and the composite self-description ranking, whereas Navran and Stauffacher (1954) report a rho of $-.03$ between these two variables. When we correlate our subjects' EPPS scores on each scale with their self-description scores, the results likewise suggest a relationship between the scales taken individually and the subjects' self-evaluation on that trait. As can be seen from Table 1, all but three of the product-moment correlations are significant at the .05 level, seven of them reaching the .01 level of significance. This is in keeping with data reported by Mann (1958) who, using more abbreviated self-descriptions, reported 10 of the 15 correlations between EPPS and self-rating to be significant at the .05 level. As we would expect from (*a*) the marked overall transparency of the EPPS items and (*b*) the relationship between EPPS scales and social desirability at the variable level, the similarity between what a subject produces on the EPPS and how he describes himself when simply asked to do so is considerable.

On one point our results coincide exactly with Navran and Stauffacher's (1954): In both studies a rho of .90 was found between composite self-report ranking and composite social desirability ranking.

As indicated earlier, we feel that the crucial questions with regard to the EPPS (or any other psychometric instrument) relate to the question of validity and not to the consideration of the elimination of the influence of test-taking attitudes. When attempts are made to evaluate the validity of the EPPS, as McKee and Turner (1961) did recently, it soon becomes clear that the lack of concern with empirical validity at the test construction stage eventuates in low predictive power. As Hathaway (1960) has suggested, it is probably better to attempt to measure and manipulate such role playing rather than to try to eliminate it. These data suggest that it is unlikely that such attempts to control for the influence of extratest factors can be largely successful. In this, they corroborate data recently published by Dicken (1959) who found that students were able to simulate various personality roles on the EPPS in spite of the social desirability controls. Our results also cast considerable doubt on the technique of selecting items on the basis of their content. It would seem that naive subjects are every bit as clever in telling us what trait a given item "measures" as are psychologists—a relationship which may not even exist in the first place!

SUMMARY

This paper addresses itself to the investigation of various extratest factors that may influence performance on the EPPS.

From the results it can be concluded that: (*a*) the EPPS items are to a large degree transparent in that their meaning can readily be guessed by naive subjects; (*b*) contrary to earlier reports, there is a significant relationship between a group's composite EPPS scores and social desirability; and (*c*) there are significant relationships between a subject's self-description and his production on the EPPS. Some implications of these findings were discussed.

REFERENCES

BARRON, F. The Edwards Personal Preference Schedule. In O. K. Buros (Ed.), *The fifth mental measurement yearbook*. Highland Park, N. J.: Gryphon, 1959. Pp. 114–117.

BROEN, W. E., JR., & WIRT, R. D. Varieties of response sets. *J. consult. Psychol.*, 1958, **22**, 237–240.

CRONBACH, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt.*, 1950, **10**, 3–31.

DICKEN, C. Simulated patterns on the Edwards Personal Preference Schedule. *J. appl. Psychol.*, 1959, **43**, 372–378.

EDWARDS, A. L. The Edwards Personal Preference Schedule. New York: Psychological Corporation, 1959.

GUSTAD, J. W. A review of the Edwards Personal Preference Schedule. *J. consult. Psychol.* 1956, **20**, 322–324.

HATHAWAY, S. R. Foreword. In W. G. Dahlstrom & G. S. Welsh (Eds.), *An MMPI handbook: A*

guide to clinical practice and research. Minneapolis: Univer. Minnesota Press, 1960.

McKEE, J. P., & TURNER, W. S. The relationship of "drive" ratings in adolescence to CPI and EPPS scores in adulthood. *Vita hum.* 1961, **4**, 1–14.

MANN, J. H. Self-ratings and the EPPS. *J. appl. Psychol.*, 1958, **42**, 267–268.

MURRAY, H. A. *Explorations in personality. A clinical and experimental study of fifty men of college age*. New York: Oxford Univer. Press, 1938.

NAVRAN, L., & STAUFFACHER, J. C. Social desirability as a factor in Edwards Personality Preference Schedule performance. *J. consult. Psychol.*, 1954, **18**, 442.

SEEMAN, W. Subtlety in structured personality tests. In G. S. Welsh & W. G. Dahlstrom (Eds.), *Basic Readings on the MMPI in psychology and medicine*. Minneapolis: Univer. Minnesota Press, 1956.

(Received July 28, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 382–383

NAYSAYING AND THE CALIFORNIA PSYCHOLOGICAL INVENTORY

DAVID ROSENHAN

Personality Research Group, Educational Testing Service, Princeton, New Jersey

Numerous investigators have noted the degree to which agreeing response sets permeate answer patterns in personality measurement (Jackson & Messick, 1958; Messick & Jackson, 1961). Moreover, the agreeing response tendency has been shown to be based on a central personality syndrome (Couch & Keniston, 1960). Thus, it may be argued that scales whose scores are substantially related to the Agreement Response Scale (Couch & Keniston, 1960) confound a number of personality variables.

This study reports the degree to which the Agreement Response Scale (ARS) is related to the scales of the California Psychological Inventory (CPI). Both tests were given to 49 undergraduates in an introductory psychology course. Significant correlations between the ARS and scales of the CPI are shown in Table 1.

TABLE 1

CORRELATIONS BETWEEN THE ARS AND CPI

Scales	Correlation with ARS ($N = 49$)
Achievement via independence	−.81***
Self-control	−.67***
Tolerance	−.64***
Psychological-mindedness	−.61***
Sense of well-being	−.58***
Responsibility	−.57***
Intellectual efficiency	−.55***
Good impression	−.54***
Capacity for status	−.46***
Achievement via conformance	−.42**
Dominance	−.32*

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Spearman rank correlations were computed between loadings on the CPI factors (Mitchell & Pierce-Jones, 1960) and the scale correlations with the ARS (Table 1). Correlations obtained were: Factor I, −.71; Factor II, .28; Factor III, −.14; Factor IV, −.22.

Of interest in these data is the pervasiveness of significant correlations; a substantial number of CPI scales, and certainly its most powerful factor, is heavily permeated with "naysaying." However, beyond noting that the disagreeing tendency accounts for a healthy amount of variance in a number of CPI scales, and that the supposedly distinct CPI personality variables are confounded, the data are not easily interpreted. Perhaps the pervasiveness of the relationship is due to the fact that "healthier" people tend to get high scores on the CPI and low scores on the ARS —although this speculation would not account for those scales in which little relationship was found with the ARS.

It is clear that the ARS is a quite adequate measure of achievement via independence. The magnitude of the relationship would sug-

gest that the latter scale measures much more of autonomy and independence (which are highly related to naysaying) than it does of achievement.

REFERENCES

COUCH, A., & KENISTON, K. Yeasayers and naysayers: Agreeing response set as a personality variable. *J. abnorm. soc. Psychol.*, 1960, **60**, 151–174.

JACKSON, D. N., & MESSICK, S. Content and style in personality assessment. *Psychol. Bull.*, 1958, **55**, 243–252.

MESSICK, S., & JACKSON, D. N. Acquiescence and the factorial interpretation of the MMPI. *Psychol. Bull.*, 1961, **58**, 299–304.

MITCHELL, J. V., & PIERCE-JONES, J. A factor analysis of Gough's California Psychology Inventory. *J. consult. Psychol.*, 1960, **24**, 453–456.

(Received August 23, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 383–386

SOCIAL DESIRABILITY AND THE RELATIVE VALIDITIES OF ACHIEVEMENT SCALES

ALFRED B. HEILBRUN, JR.

University of Iowa

Certainly one of the major recent issues in objective psychometrics has been the effect of response sets or styles upon measurement and prediction. Although primary credit is typically given to Cronbach (1946, 1950) for his early descriptions of various response sets and their presumed effects upon test scores, the later work of Edwards with the social desirability variable and its importance with *Q*-sort measures (1955), his own Personal Preference Schedule (EPPS) (1953), and the MMPI (in press) has proven to be a major research stimulus to those interested in the antecedents of performance on objective personality questionnaires. However, from the point of view of the present writer, there has been an unfortunate over-concentration in social desirability (*SD*) research upon whether social values are related to various test performances. Although these relationships are of interest, they fail to provide answers to the crucial questions of how the validity of a test is affected by *SD* and whether these effects vary as a function of the test variable measured or the criterion predicted. For example, to find that the Edwards' *SD* scale (1957) is correlated −.70 with a predictive scale like the MMPI *D* scale (Fordyce, 1956) does not allow one to assume that prediction from the *D* scale is influenced unless there is independent evidence that the *SD* variable is a suppressor variable (i.e., related to the predictor variable and unrelated to the criterion variable). It is true that one can argue, as Edwards does with some cogency (Edwards, Diers, & Walker, 1962), that a scale designed to measure a specified trait but which includes items keyed for socially desirable responses can be considered another form of the *SD* scale. However, the crucial factor in determining what a presumed trait scale which correlates highly with *SD* actually measures remains its relationships with acceptable nontest criteria for that trait.

Heilbrun and Goodstein (1961b) have assessed the extent to which the tendency to respond to the social desirability of scale items influenced the predictive validity of two EPPS scales—Achievement and Abasement. By obtaining an independent measure of the *SD* values imposed upon EPPS statements by each college subject, individually "tailored" scales were developed in which the pairs of item statements were as perfectly matched for *SD* value for each individual as possible. Predictions from these revised achievement and abasement scales were compared with similar predictions from the standard scales for the same subjects. If response to differential *SD* values of statements on the EPPS was a source of error, then the revised scales should have proven more valid. However, the results showed the revised scales to be less predictive than the standard scales. At the

least, these results suggest that considerable caution should be taken in continuing the unqualified assumption that reducing the contingency of personality scale scores upon *SD* will result in increased validity. At most, such results suggest that scale performance variance attributable to *SD* is true rather than error variance in some predictive situations.

The present study represents a further assessment of the nature of *SD* variance on personality questionnaires. All other things being equal, if response to *SD* contributes to test invalidity, then a scale which measures a given trait and reduces the role of *SD* in performance should be more valid than a scale, measuring the same variable, for which *SD*'s effect is left uncontrolled in the extent it can influence performance.

METHOD

Achievement Scales

The *SD*-controlled Achievement scale of the EPPS was developed rationally and includes 28 items, each incorporating a need achieving statement and a second statement measuring a different need variable. The testee is asked to select the more characteristic or the better liked statement in each pair and the frequency of selection of achievement need statements is presumed to reflect the relative strength of that need within the person's need hierarchy. These statement pairs were originally matched for *SD* value and although there is evidence that subjects still tend to endorse higher *SD*-valued behaviors (Heilbrun & Goodstein, 1961a), the same study showed that about 29% of the EPPS items provide perfect *SD*-matched statement pairs on the average. Thus, on many EPPS items the subject is presumably left without the option of selecting an alternative based on higher *SD* value.

Included in the Need Scales (NS) developed by Heilbrun (1959) is an *SD*-uncontrolled Achievement scale measuring the same variable as measured by the EPPS Achievement scale. The NS Achievement scale is scored from the Adjective Check List (ACL) (Gough, 1960) by subtracting the number of adjectives endorsed as self-characteristic which contraindicate a high achievement need from the number of endorsed adjectives indicating the presence of high n achievement in the endorser. There are no built-in checks for *SD* on the ACL such as forced-choice items or subtle items and except for instructional encouragement to be frank the testee is unrestricted in his endorsement of adjectives as far as their *SD* is concerned. This is clearly portrayed by a previous finding (Heilbrun, 1958) that the percentage of adjectives endorsed within an earlier form of the Need Scales correlated .98 with the saturation of those scales with adjectives judged to reflect favorably on the endorser.

Criterion and Subjects

College academic performance was selected as the criterion for assessing the relative validities of the *SD*-controlled EPPS Achievement scale and the *SD*-uncontrolled NS Achievement scale. As a preliminary step, achieving and nonachieving academic performance was defined by plotting undergraduate college grades against an ability measure, the university entrance examination composite percentile, for males ($N = 149$) and for females ($N = 127$) separately. A best fit line was drawn for each plot (a straight-line proved to be the best fit in both cases); thus, a cumulative grade-point average which fell above the line defined an achieving student whereas an average which fell below defined the nonachiever. It should be noted that this procedure serves to partial out intellectual ability in assessing the relationships between the need achievement scales and actual college achievement.

The subjects for the present study were 156 undergraduates from a large psychology class at the University of Iowa, including 76 females and 80 males. Plotting their entrance composite percentiles against their cumulative grade-point averages resulted in 67 achievers (25 males and 42 females) and 89 nonachievers (55 males and 34 females). These subjects had been group-administered the ACL and the EPPS in that order as part of a 2-hour testing battery.

TABLE 1

COMPARISON OF PREDICTIVE VALIDITY OF EPPS AND NS NEED–ACHIEVEMENT SCALES

Need achievement scale	Male college achievers ($N = 25$)		Male college nonachievers ($N = 55$)			Female college achievers ($N = 42$)		Female college nonachievers ($N = 34$)		
	M	*SD*	*M*	*SD*	*t*	*M*	*SD*	*M*	*SD*	*t*
EPPS (*SD*-controlled)	54.08	10.49	52.62	7.85	.69	55.21	10.23	54.56	10.52	.27
HS (*SD*-uncontrolled)	54.72	10.42	48.15	11.40	2.43**	50.21	8.64	48.50	10.84	.76

** Denotes significance beyond .01 level of confidence (one-tailed).

RESULTS AND DISCUSSION

The mean scores on the EPPS Achievement scale and the NS Achievement scale for the male and female achieving and nonachieving groups were compared and the results are given in Table 1. The *SD*-controlled EPPS scale failed to discriminate between either the male or the female achievers and nonachievers although the differences in means were in the appropriate direction. The *SD*-uncontrolled NS scale did reliably distinguish between the male achievers and nonachievers in the predicted direction. The female achievers were higher on the NS Achievement scale but not significantly so.

A more direct comparison between the two achievement scales and one which more clearly portrays their relative validities was accomplished by determining the number of "hits" and "misses" afforded by classifying all subjects with achievement scores of $T = 50$ or below as nonachievers and all subjects with scores of $T > 50$ as achievers and comparing these classifications with the actual achievement status of the subjects. By this system the EPPS Achievement scale appropriately classified 32 male subjects and misclassified 48 male subjects. The NS Achievement scale correctly classified 46 male subjects and missed on 34 male subjects. A chi square test provided a value of 4.23 for 1 *df* which is significant at the 5% level; thus, the *SD*-uncontrolled NS scale more accurately classified male achievers and nonachievers than did the *SD*-controlled EPPS scale. For females, the EPPS scale correctly classified 41 subjects and missed on 35 subjects, and the NS scale hit on 48 subjects and missed on 28 subjects. The resulting χ^2 value of .98 for 1 *df* was not significant ($p > .30$); thus, there was no difference in the discriminatory ability of the *SD*-controlled and *SD*-uncontrolled scales for females.

The results of this study clearly fail to support the contention that curtailment of the effect of social desirability increases the validity of an objective personality scale. In fact, the evidence indicates that the opposite is true in the prediction of college achievement for males. The NS Achievement scale which does not limit the effect of *SD* proved to be more valid in its relationship to academic performance than was the EPPS *SD*-controlled counterpart.

Although it is true that the differences in *SD*-control in the present study are confounded with differences in the manner in which need achievement is measured (e.g., statement endorsement vs. adjective endorsement), it may be recalled that Heilbrun and Goodstein (1961b) found that increased curtailment of *SD* as a basis for statement endorsement tended to lower the validity of EPPS scales when the formal characteristics of the scales were held constant. These results suggest to the present investigator that the major research focus should be shifted from establishing relationships between *SD* and personality scales to how *SD* influences the relationships between scales and their criteria. Far too little cognizance has been taken of the possibility that the social value imposed upon questionnaire alternatives is not a suppressor variable but, at least under certain conditions, may contribute to predictive validity.

SUMMARY

The present study compared the Achievement scale scores taken from the Edwards Personal Preference Schedule and the Heilbrun Need Scales for male and female groups of college students defined as achievers and nonachievers depending upon whether their grade-point averages fell above or below a prediction based upon intellectual ability. The hypothesis tested was that responding to the social desirability (*SD*) of personality item alternatives is a source of test invalidity. If true, the EPPS scale in which *SD* is controlled to some extent should be more valid than the NS scale which places no restrictions on the contribution of *SD* to item responses. The results failed to support this hypothesis, since the *SD*-uncontrolled NS Achievement scale proved to be more valid than the EPPS scale in predicting male achievement and no between-scale difference was found for females. Suggestions regarding the direction of future *SD* research were made.

REFERENCES

CRONBACH, L. J. Response sets and test validity. *Educ. psychol. Measmt.*, 1946, **6**, 475–494.

CRONBACH, L. J. Further evidence on response sets and test design. *Educ. psychol. Measmt.*, 1950, **10**, 3–31.

EDWARDS, A. L. The relationship between judged desirability of a trait and the probability that the trait will be endorsed. *J. appl. Psychol.*, 1953, **37**, 90–93.

EDWARDS, A. L. Social desirability and *Q* sorts. *J. consult. Psychol.*, 1955, **19**, 462.

EDWARDS, A. L. *The social desirability variable in personality assessment and research.* New York: Dryden, 1957.

EDWARDS, A. L. Social desirability and expected means on MMPI scales. *Educ. psychol. Measmt.*, in press.

EDWARDS, A. L., DIERS, CAROL, & WALKER, J. N. Response sets and factor loadings on 61 personality scales. *J. appl. Psychol.*, 1962, **46**, 220–225.

FORDYCE, W. E. Social desirability in the MMPI. *J. consult. Psychol.*, 1956, **20**, 171–175.

GOUGH, H. G. The Adjective Check List as a personality assessment research technique. *Psychol. Rep.*, 1960, **7**, 107–122.

HEILBRUN, A. B. Relationships between the Adjective Check List, Personal Preference Schedule and desirability factors under varying defensiveness conditions. *J. clin. Psychol.*, 1958, **14**, 283–287.

HEILBRUN, A. B. Validation of a need scaling technique for the Adjective Check List. *J. consult. Psychol.*, 1959, **23**, 347–351.

HEILBRUN, A. B., & GOODSTEIN, L. D. The relationships between individually defined and group defined social desirability and performance on the Edwards Personal Preference Schedule. *J. consult. Psychol.*, 1961, **25**, 200–204. (a)

HEILBRUN, A. B., & GOODSTEIN, L. D. Social desirability response set: Error or predictor variable? *J. Psychol.*, 1961, **51**, 321–329. (b)

(Received October 11, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 386–388

BENDER–GESTALT FIGURE ROTATIONS IN THE MENTALLY RETARDED

A. B. SILVERSTEIN AND PHILIP J. MOHAN

Pacific State Hospital

A recent paper by Griffith and Taylor (1960) provided statistics on the incidence of Bender-Gestalt figure rotations in a psychiatric hospital. The authors reported that mentally retarded patients had a significantly higher incidence of rotations than those with chronic brain syndrome, and that both of these groups had a significantly higher incidence than neurotics, schizophrenics, and patients with character disorders. Because of the relatively small number of retardates ($N = 59$), their incidence of rotations was not considered to be as reliable as those for the other diagnostic groups. It was also noted that in a psychiatric hospital, the diagnosis of mental retardation is always suspect. Nevertheless, the negative relationship between rotations and intelligence was supported by a significant difference between the IQs of nonretarded patients who had at least one rotation, and those who had none.

The aims of the present study were as follows: (*a*) to obtain statistics on the incidence of Bender rotations in a hospital for the mentally retarded; (*b*) to assess the stability of rotations over time; and (*c*) to test the relationships between rotations and the variables of age, sex, IQ, diagnosis, and hospitalization.

PHASE I

Method

The Bender is not part of a standard battery of psychological tests at Pacific State Hospital. Psychologists administer the test at their discretion, primarily to patients in whom brain damage or psychosis is suspected. A search of the files of the Psychology Department revealed 1,344 patients who had been given the test, 281 of them more than once. In order to assess the stability of rotations over time, it was decided to focus first on those patients who had been retested. However, not all of these patients could be included in the sample, either because their test records could not be located, because the test had not been given in its entirety, or because the patients had been diagnosed as not mentally retarded.

The final sample consisted of 200 patients, 130 males and 70 females. The initial test had been given to 97 of these patients during the preadmission period, i.e., before they actually entered the

hospital. The median age of the sample at the time of the initial test was 16.0, and at the time of the retest, 18.9; thus the average test-retest interval was almost 3 years. During this interval, the median IQ of the sample changed only slightly, from 60.6 to 61.2. For 155 of the patients, there was no clinical or historical indication of organic disease or pathology.

Previous investigators have used various criteria for scoring rotations. In the present study, the definition used by Griffith and Taylor was adopted: "an angular displacement of 45 degrees or more in a figure which otherwise remained recognizable" (p. 189). Almost all of the test records could be scored by inspection; a protractor was used for those figures where the scoring was in doubt.

Results and Discussion

Forty-six percent of the patients had at least one rotation on the initial test. This value did not differ significantly at the .05 level from that for the retest, 52%, nor from that reported by Griffith and Taylor for retardates in a psychiatric hospital, 55.9%. The incidence of rotations for the individual figures ranged from 0.5% for Figure 6 to 22.5% for Figure 3.

The relationship between rotations on the two tests, as measured by the phi coefficient, was .26, very low but highly significant ($\chi^2 = 12.92$, $p < .001$). Phi coefficients for the individual figures ranged from .00 for Figure 6 to .32 for Figure 3. Even allowing for the length of the average test-retest interval, the stability of rotations was not impressive.

Relationships between rotations and age, sex, IQ, diagnosis (organic versus nonorganic), and hospitalization were tested by means of chi square, the continuous variables being dichotomized as close to the median as possible. On the initial test, rotations were not significantly related to any of these variables. On the retest, 59.8% of the patients with IQs below 62 had rotations, compared to 43.9% of those with IQs of 62 and above ($\chi^2 = 4.46$, $p < .05$). The negative relationship between rotations and intelligence reported by Griffith and Taylor for nonretarded psychiatric patients was thus confirmed for retardates, though only on the retest.

Also on the retest, 65.7% of the female patients had rotations, compared to 44.6% of the male patients ($\chi^2 = 7.29$, $p < .01$). This significant sex difference was unanticipated and cannot readily be accounted for. A partial explanation appears to lie in the following interrelated findings: (*a*) the IQs of the male patients were significantly higher than those of the female patients at the time of the retest ($\chi^2 = 5.28$, $p < .05$), but not at the time of the initial test; (*b*) the IQs of the male patients increased significantly during the test-retest interval ($\chi^2 = 22.27$, $p < .001$), whereas those of the female patients did not; and (*c*) the incidence of rotations for the female patients increased appreciably from the initial test to the retest ($\chi^2 = 3.38$, $p < .07$), whereas that for the male patients did not. These interactions involving sex are themselves rather puzzling. The possibility that they were due to a significant sex difference in the test-retest interval was explored, but this proved not to be the case.

As for the other findings, the failure to obtain significant diagnostic differences was somewhat disappointing but scarcely surprising. The literature on mental retardation contains many studies comparing exogenous and endogenous, brain injured and familial, organic and nonorganic patients, and in general, the differences between such groups are small in number and inconsistent from study to study.

PHASE II

Method

Lingering doubts about the representativeness of the Phase I sample, reinforced by the editor of this journal, led to the decision to draw a second sample of 200 patients, from the 1,063 who had been tested with the Bender only once. The selection of patients was at random, subject only to the restrictions that their test records could be located, that the test had been given in its entirety, and that the patients had been diagnosed as mentally retarded.

The final sample consisted of 118 males and 82 females. The test had been given to 53 of these patients during the preadmission period. The median age of the sample was 18.8, and their median IQ was 56.6. For 133 of the patients, there was no indication of organicity. When this sample was compared with the Phase I sample on the variables of age, sex, IQ, diagnosis, and hospitalization, the two were found to differ significantly on every variable but sex. The Phase II sample was significantly older ($\chi^2 = 14.44$, $p < .001$) and had a significantly lower IQ ($\chi^2 = 4.00$, $p < .05$) than the Phase I sample. A significantly greater proportion of the Phase II sample was organic ($\chi^2 = 5.47$, $p < .05$), and a significantly smaller proportion had been tested dur-

ing the preadmission period ($\chi^2 = 19.72$, $p < .001$). This last difference is attributable to the fact that a significantly greater proportion of the Phase I sample came to the hospital after the establishment of a preadmission service in 1952. The remaining differences appear to result from the policy of the Psychology Department in regard to retesting. All patients are routinely retested at least once every 5 years, but patients may be retested after a shorter interval, if a change in their clinical status has been observed, and/or a modification in their treatment is being contemplated. On both of these counts, it is understandable that the patients who had been retested were younger and brighter than those who had been tested only once, and that fewer of them were organic.

Rotations were scored using the same definition that had been adopted in Phase I.

Results and Discussion

Thirty-nine percent of the patients had at least one rotation on the test. This value did not differ significantly from that for the initial test in Phase I, 46%, nor from that reported by Griffith and Taylor, 55.9%. The incidence of rotations for individual figures was generally similar to that in Phase I, ranging from none for Figure 6 to 17.5% for Figure 3.

Again, relationships between rotations and the five background variables were tested by means of chi square, and as on the initial test in Phase I, none of these relationships was significant. In short, the differences between the samples in the two phases did not markedly influence the results obtained. Approximately 40 to 50% of the patients had one or more rotations on the test, and (except for the curious results on the retest in Phase I), rotations were not significantly related to age, sex, IQ, diagnosis, or hospitalization.

SUMMARY AND CONCLUSIONS

The Bender-Gestalt records of 400 patients in a hospital for the mentally retarded were scored for figure rotations. The Phase I sample consisted of 200 patients who had been given the test more than once, while the 200 patients who comprised the Phase II sample had been tested only once. The two samples differed significantly in age, IQ, diagnosis (organic versus nonorganic), and hospitalization, but these differences did not markedly influence the results in the two phases.

Approximately 40 to 50% of the patients had at least one rotation on the test, a finding consistent with that reported previously for retardates in a psychiatric hospital. The stability of rotations over an average test-retest interval of almost 3 years was significant but low. Rotations were significantly related to sex and IQ on the retest in Phase I, but not on the initial test, and not in Phase II. In neither phase were rotations significantly related to age, diagnosis, or hospitalization.

REFERENCE

GRIFFITH, R. M., & TAYLOR, VIVIAN H. Incidence of Bender-Gestalt figure rotations. *J. consult. Psychol.*, 1960, **24**, 189–190.

(Received October 13, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 389

BRIEF REPORTS

OBSERVABLE SIGNS OF ANXIETY OR DISTRESS DURING PSYCHOTHERAPY [1]

ALLEN RASKIN

Veterans Administration, Washington, D. C.

This study investigated the frequency of occurrence of 15 observable anxiety signs during initial therapy interviews and their relative importance to the therapist as indicators of patient anxiety or distress.

Data were collected in 23 Veterans Administration Mental Hygiene Clinics. Study patients were all males, under Age 50, who were acceptable for individual psychotherapy scheduled for 50 minutes once a week.

At the conclusion of the patient's first psychotherapy interview, a list of 15 behaviors or symptoms commonly associated with anxiety was provided with instructions to the therapist to circle each symptom, sign, or act observed during the initial interview. Following this interview, therapists also rated patients on four-point graphic scales of overt tension, extent of verbal expressions of being concerned or apprehensive during the therapy hour, extent of verbal expressions of being concerned or apprehensive in situations outside therapy, and degree to which the patient is inwardly distressed, agitated or otherwise suffering from his symptoms and problems.

Use of a scaling frame (duMas, 1955) permitted visual indentification of the anxiety signs whose presence was closely associated with above-median ratings and whose absence was associated with below-median ratings on each of the four graphic rating scales. Forty-eight cases, randomly selected from a pool of 234 cases, were plotted on the scaling frame. The remaining cases were randomly divided into two cross-validation samples of 93 cases each.

The median number of anxiety signs reported for all patients was 2.44. The three signs most frequently observed by therapists were: (*a*) body musculature appears taut, strained or tense, (*b*) moves and shifts restlessly, and (*c*) uses hands excessively. These three and one additional sign, exhibited tics or twitching of the face or neck, were the best predictors of therapist ratings of overt tension, verbalized concern or anxiety in situations outside therapy, and inner distress and discomfort. The presence of at least two of these signs was significantly associated with above-median scores, and conversely, the presence of less than two of these signs was associated with below-median scores on these three rating scales. This finding was consistent for the two cross-validation samples as well as the original sample. It is of interest to note that these anxiety signs all describe excessive or unusual motor behavior. Breathing difficulties, vascular disturbances, such as sweating and face flushing, and speech disturbances, were less closely associated with other judgments of patient anxiety and distress.

None of the anxiety signs, taken singly or in combination, were significantly associated with therapist ratings of verbalized concern or anxiety during therapy. However, the data indicated that many patients rated as distinctly or conspicuously tense, evidencing much or extreme distress or discomfort, and who displayed at least two of the aforementioned anxiety signs, did not report feeling unduly tense or anxious during therapy.

[1] An extended report of this study may be obtained without charge from Allen Raskin (Research Department, Department of Psychiatry, D. C. General Hospital, Washington 3, D. C.) or for a fee from the American Documentation Institute. Order Document No. 7186 from ADI Auxilliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

REFERENCE

duMas, F. M. *Manifest structure analysis.* Missoula: Montana Univer. Press, 1955.

(Received June 23, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 390

A FURTHER INVESTIGATION OF THE RELATION BETWEEN ANXIETY AND DIGIT SYMBOL PERFORMANCE [1]

ROBERT A. JOHNSTON AND HERBERT J. CROSS

University of Richmond

The major purpose of the present study was to further investigate the relationship between anxiety as measured by the Taylor Manifest Anxiety (*MA*) scale and performance on the Digit Symbol Test (DST) by manipulating the difficulty level of the criterion task. If we accept a drive interpretation of anxiety (A), high A subjects should be superior to low A subjects when the DST is made easy. The relative performances should reverse when the task is made more difficult. A second purpose was to investigate the DST itself, looking at its relationship to learning ability and speed of copying.

The *MA* scale and a copying task were administered to 109 general psychology students. School and College Ability Test (SCAT) scores were also available on each subject. The subjects were rank ordered on *MA* scale score and the 30 highest and 30 lowest scorers were randomly assigned to three treatment groups. The easy task group was taught the correct symbols for the digits and the difficult task group was taught incorrect (jumbled) symbols. The control group was taught to associate letters and pictures of objects. All learned by a rote paired associates method to six successive errorless recalls, after which two written practice trials were given.

An analysis of variance showed a significant effect of difficulty level, suggesting that the experimental procedures had been effective in making the task easy, although the differences between the control and difficult task group was not significant ($p = .05$). Neither the main effect of anxiety, nor the anxiety by difficulty interaction was significant, however, suggesting that anxiety was not related to performance on the DST. The correlation between SCAT total score and DST was .35 ($p = .05$). An analysis of covariance using SCAT scores as a control factor yielded the same results. A .23 correlation ($p = .07$) was found between copying speed and DST scores. Copying speed scores were used as a second control factor in performing a multiple analysis of covariance. Again, difficulty level was the only significant factor. An analysis of variance of copying speed showed that high A subjects copy more rapidly than do low A subjects ($p = .025$).

The lack of a relationship between anxiety as measured by the *MA* scale and performance on the DST was confirmed. There is no evidence that this lack of relationship is due to difficulty level of the DST. It could be argued that the procedure was not successful in making the task difficult; however, there is no similar explanation for the lack of superiority of the high A group when the task was made demonstrably easy.

The superiority of high A subjects on copying speed would suggest that the *MA* scale does differentiate drive levels. The lack of a relationship between *MA* scale and DST could be caused by a small proportion of DST variance attributable to speed factors, and thus associated with drive.

Also, it has been suggested that high A subjects might have a disadvantage due to the interference of anxiety which, coupled with an advantage due to speed, could obscure any trends due to anxiety per se.

(Received September 1, 1961)

[1] An extended report of this study may be obtained without charge from Robert A. Johnston (University of Richmond, Department of Psychology, Richmond, Virginia) or for a fee from the American Documentation Institute. Order Document No. 7187 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 391

A NOTE ON THE VALIDITY OF THE WECHSLER ADULT INTELLIGENCE SCALE FOR MENTAL RETARDATES[1]

GARY M. FISHER

Fairview State Hospital, Costa Mesa, California

It has been shown (Fisher, Kilman, & Shotwell, 1961) that, among retardates, the WAIS gives significantly higher IQs than the Stanford-Binet (S-B), L-M, and that the discrepancy between the two IQs is affected by age. The purpose of this paper is to present data which suggest that the WAIS overestimates the IQ of mental retardates.

The subjects in this study were 76 endogenous retardates divided into four groups on the basis of age at initial S-B ($S\text{-}B_1$). Group A had the S-B between the ages of 6–14 years, and Groups B, C, and D between the ages of 15–34 years. A second S-B, Form L-M ($S\text{-}B_2$) and WAIS were administered (within a year of one another) when the subjects were between the ages of 15–34, 35–44, 45–54, and 55–70 years, respectively for Groups A, B, C, and D. The mean test-retest time interval was 20 years for Groups A and B, and 26 years for Groups C and D. For the four groups, the following mean IQs were obtained: $S\text{-}B_1$: 47.00, 47.72, 47.60, and 48.17; $S\text{-}B_2$: 46.79,[2] 50.00, 50.56, and 45.33; WAIS: 55.36, 59.28, 62.84, and 62.50. The correlations between $S\text{-}B_1$ and $S\text{-}B_2$ IQs for the four groups were .65, .91, .87, and .78. The differences in means for the two S-B IQs were 0.18, 2.28, 2.96, and 2.84 IQ points, the difference being significant ($p < .05$) for Groups B and C. The correlations between $S\text{-}B_1$ and WAIS IQs for the four groups were .32, .81, .74, and .70. WAIS IQ was significantly higher ($p < .01$) than $S\text{-}B_1$ IQ by 8.36, 11.56, 14.88, and 14.33 IQ points, respectively. Correlations between age at $S\text{-}B_1$ with $S\text{-}B_1$ IQ (.04), age at retest with $S\text{-}B_2$ IQ (−.06), and age at retest with WAIS IQ (.00) were all insignificant, whereas age at retest correlated .37 ($p < .01$) with WAIS minus $S\text{-}B_2$ IQ.

Three factors of importance emerge from this study: (*a*) for subjects tested with the S-B in their formative years (6–14) and later retested at the peak of their intellectual efficiency (15–34 years), with S-B and WAIS, there is constancy of S-B IQs, whereas WAIS IQs are significantly higher than $S\text{-}B_1$ IQs; (*b*) subjects tested with the S-B at their peak years of intellectual efficiency give fairly constant S-B IQs when retested during later years (35–70), whereas their later WAIS IQs are all significantly higher; (*c*) WAIS IQs become increasingly larger than S-B IQs as age increases.

The data suggest that for mental retardates, the S-B gives a reliable measure of intelligence over a long period of time and that IQs from the WAIS overestimate the reliable S-B IQs, are artifactually affected by age and, therefore, have questionable validity. It is suggested that the norms for the WAIS may be inadequate in that there was insufficient sampling of subjects of lower intellect and, as a result, the rate of decline of intellectual abilities with age for the standardization sample is inapplicable to mental retardates.

[1] An extended report of this study may be obtained without charge from G. M. Fisher (Fairview State Hospital, Costa Mesa, California) or for a fee from the American Documentation Institute. Order Document No. 7188 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

This study was supported, in part, by a grant from the California State Department of Mental Hygiene (#R-58-14.1-3).

[2] IQs from the S-B, Form L-M, were converted from deviation to ratio IQs to make them comparable with IQs from earlier forms of the S-B.

REFERENCE

FISHER, G. M., KILMAN, BEVERLY A., & SHOTWELL, ANNA M. Comparability of intelligence quotients of mental defectives on the Wechsler Adult Intelligence Scale and the 1960 revision of the Stanford-Binet. *J. consult. Psychol.*, 1961, **25**, 192–195.

(Received September 7, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 392

THE EFFECT OF REPEATED TESTING ON ANXIETY SCALE SCORES [1]

A. W. BENDIG AND GAIL BRUDER

University of Pittsburgh

Scheier and Cattell (1960) have published a battery of eight 50-item questionnaires measuring Cattell's second-order "anxiety" factor. Since only a few items overlap each pair of scales, the battery appears useful in repeatedly measuring anxiety level in experimental research. However, Windle (1954, 1955) found that subjects appear to become slightly more "normal" or less anxious on the second administration of a personality questionnaire. In the studies reviewed by Windle only two administrations were used and the same items were used on both administrations. It is not known whether a similar decrease would occur when different items are used or whether the decrease in scores would continue on further administrations.

The eight Scheier-Cattell scales were randomly divided into the subbatteries and assigned to the cells of a $4 \times 4 \times 4$ balanced latin square design. Each of the two squares was replicated twice with the subjects of one square having 2 days between administrations of the forms and the subjects of the second square having 7 days between administrations. The 48 subjects were college students recruited from an introductory psychology course and they were randomly assigned to the rows of each square. The 12 subjects assigned to a given square (three to each order) were given the four forms of one of the subbatteries with either 2 or 7 days intervening between each form. The trichotomous item scores (0, 1, or 2) were summed over all 50 items on each form.

Two analyses of variance (one for each subbattery) of the replicated Latin square design indicated that neither the order of administration of the forms nor the time intervals between administrations had any significant effect on the anxiety scores. The order by time interaction was significant for one subbattery, but not for the other. The differences between administrations and the interaction of time intervals and administrations did not approach significance for either subbattery and inspection of the means gave no evidence of consistent changes in scores between administrations. The absence of trend effects was apparent for both subbatteries and for both time intervals. The differences among forms within a subbattery were significant for both subbatteries.

The time interval (2 or 7 days) had no effect on the interform equivalents reliabilities (average correlation with the other forms in the subbattery) of the questionnaires and these reliabilities (.60 to .85) are somewhat larger than those reported by Scheier and Cattell (1960, p. 2).

The tendency for questionnaire scores to become more normal appears attributable to using the same items or forms at each administration. No statistically significant or consistent changes occurred between the first and second administrations, nor between the later administrations. The absence of systematic change was confirmed for two different subbatteries of parallel forms and for two different time intervals between administrations.

[1] An extended report of this study can be obtained without charge from A. W. Bendig (Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania) or for a fee from the American Documentation Institute. Order Document No. 7189 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $1.25 for microfilm and $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

REFERENCES

SCHEIER, I. H., & CATTELL, R. B. *Temporary handbook for the IPAT 8-parallel form anxiety battery.* Champaign, Ill.: Institute for Personality and Ability Testing, 1960.

WINDLE, C. Test-retest effect on personality questionnaire. *Educ. psychol. Measmt.*, 1954, **14**, 617–633.

WINDLE, C. Further studies of test-retest effect on personality questionnaires. *Educ. psychol. Measmt.*, 1955, **15**, 246–253.

(Received September 21, 1961)

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 393

RACE OF EXAMINER AND PATIENT IN PSYCHIATRIC DIAGNOSIS AND RECOMMENDATIONS [1]

ELAINE DORFMAN AND ROBERT J. KLEINER

Eastern Mental Health Center, Philadelphia

Valid race comparisons on mental illness rates require equivalent diagnostic study of each group. Most psychiatrists being white, Negroes are usually evaluated in a cross-racial interview, a condition rarely true for white examinees. Possible effects of this difference can be studied where white and Negro patients are randomly assigned for diagnosis by psychiatrists of both races.

Case records of first admissions to a state reception center were searched to select patients aged 20 through 59 years, known to be white or Negro, and diagnosed by one of five psychiatrists (three white and two Negro). Two time periods were used, one of 18 months ($N = 1675$), the second of 4 months ($N = 494$). Chi-square tests showed no race difference among caseloads of the five doctors. Thus, the requirement of unbiased assignment was met for the race variable during both periods.

From the 1675 records, a sample of 940 cases was drawn to include 94 white and 94 Negro cases per doctor, in a mixed factorial and repeated measurements design. It was factorial in that the variables, race of patient and examiner, were studied in all possible combinations. But since the same psychiatrists examined both the white and Negro patients, there were also repeated measurements of the work of these doctor-subjects. Therefore, two error terms were used in variance analyses: one for the independent data comparing white and Negro doctors, and one for the possibly correlated data from the same doctors.

Within each group of 94 cases of each race per doctor, frequency counts of various psychiatric diagnoses were the raw data for analyses of variance. Frequencies of schizophrenia, neurosis, personality disorder, organic brain syndrome, and recommended hospitalization were separately analyzed. Other categories occurred too rarely for analysis by race of patient and examiner. Results of appropriate F tests showed no effects of race of patient, race of examiner, or the interaction between the two.

The second set of 494 records was used to check the above findings. A sample of 270 cases, with 27 white and 27 Negro patients per doctor, was drawn. Only schizophrenia, neurosis, and recommended hospitalization occurred often enough for analysis of variance, so that the previous findings on personality disorder and organic brain syndrome could not be checked. As in the first sample, race proved irrelevant to the two diagnoses. For recommended hospitalization, the patient's race and the interaction between race of patient and examiner were again insignificant. In contrast, Negro doctors recommended hospitalization less often than white doctors ($p < .01$). The last is viewed as a chance effect, since it is the only 1 of 24 significance tests to reach an acceptable level, occurred only in the smaller sample, and had not been specifically predicted.

It seems fair to conclude that within the limits of our conditions, race is irrelevant to diagnosis and recommendations. Possible generalization to other diagnoses, recommendations, and psychiatrists must await replicated studies in other settings. Since we used two patient samples from the same population, the findings can probably be generalized to the work of the same doctors with other patients in this setting.

(Received October 1, 1961)

[1] An extended report of this study may be obtained without charge from Elaine Dorfman (2314 Delancey Place, Philadelphia 3, Pennsylvania) or for a fee from the American Documentation Institute. Order Document No. 7191 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

Journal of Consulting Psychology
1962, Vol. 26, No. 4, 394

PERSONALITY CHARACTERISTICS (EPPS), LEVEL OF EXPECTATION, AND PERFORMANCE [1]

CARROLL E. IZARD

Vanderbilt University

The present study aimed to relate selected scales of Edwards' Personal Preference Schedule (EPPS) to level of expectation and actual performance. Level of expectation is derived from the subject's estimate of what he will actually score on a future performance. This measure is related to, but not synonymous with, level of aspiration.

The performance and level of expectation measures were functions of a real life situation. The performance was part of the subject's ongoing responsibility. The level of expectation index was derived from the brief, probably ego-peripheral exercise of estimating future performance. In view of this, it was hypothesized that EPPS Achievement would correlate significantly with actual performance but not with level of expectation. A general prediction was that more of the remaining personality characteristics (EPPS variables) would correlate significantly with performance than with level of expectation. The author's students in a personality course, 53 females and 33 males, took the EPPS at the beginning of the semester. Level of expectation was computed as the subject's estimated score for an objective course examination that was described as comparable in length and difficulty to one on which he had recently had feedback. The index of performance was the number of items correct on all examinations in the course.

Of the 30 possible correlations between EPPS variables and level of expectation, only Endurance for the male group was significant. Several of the correlations between EPPS variables and actual performance were significant: for males, Achievement (.40, $p < .05$) and Abasement ($-.70$, $p < .01$); for females, Achievement, Dominance, and Change (.33, .33, $-.32$, respectively, $p < .02$) and Nurturance ($-.48$, $p < .01$).

The difference between these data and that of Atkinson (1957) regarding Achievement and level of expectation may be accounted for in terms of differences both in measuring devices and testing conditions. The discrepancy between the present results and Atkinson and Litwin's (1960) finding of a nonsignificant relationship between EPPS Achievement and final examination score is not so readily explained. Data from a different sample supported the present finding of a significant relationship between these variables. EPPS Achievement scores obtained on entering college were correlated with cumulative grade point average in the senior year: for males, $N = 180$, $r = .28$, $p < .01$; for females, $N = 98$, $r = .17$, $.05 < p < .10$. This strongly supports the hypothesis of a real, though low, relationship between EPPS Achievement and academic achievement among male subjects.

The fact that EPPS Achievement correlated significantly with performance on a task that was part of the real life situation for the subjects involved substantiates the validity of this scale as a predictor of at least one type of actual, meaningful achievement, particularly for male subjects.

[1] This study was supported by Vanderbilt-ONR Contract Nonr 2149(03), C. E. Izard, principal investigator. The opinions and conclusions do not necessarily reflect those of the United States Department of the Navy.

An extended report of this study may be obtained without charge from Carroll E. Izard (Vanderbilt University, Nashville, Tennessee) or for a fee from the American Documentation Institute. Order Document No. 7192 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

REFERENCES

ATKINSON, J. W. Motivational determinants of risk-taking behavior. *Psychol. Rev.*, 1957, **64**, 359–372.

ATKINSON, J. W., & LITWIN, G. H. Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *J. abnorm. soc. Psychol.*, 1960, **60**, 52–63.

(Received November 14, 1961)

*Four of a series

Therapeutic Education

At Devereux Schools Education is a therapeutic process.

Academic performance is regarded as only one function of the total personality. Competition is structured and controlled according to the child's emotional maturity and intellectual capabilities.

Curriculum ranges through all the offerings of the usual public school system, including business and vocational courses. Remedial work in reading and speech, tutoring for specific learning disabilities and other special services are also available and incorporated in the child's program if indicated.

Instruction is highly individualized, consisting of private lessons or work done in very small groups.

For Information and Literature

Devon, Penna.:
Charles J. Fowler, *Director of Admissions*
J. Clifford Scott, M.D., *Director of Psychiatry*
John R. Kleiser, Ph.D., *Director of Clinical Psychology*

Santa Barbara (Box 1670), Calif.:
Keith A. Seaton, *Registrar*
Robert G. Ferguson, Ed.D., *Director of Education and Training*
Kenneth L. Grevatt, M.D., *Director of Medical Services*

Victoria (Box 2200), Texas:
Kenneth K. Knopp, *Registrar*
Richard G. Danko, *Chief Administrator*
George M. Constant, M.D., *Consulting Psychiatrist*

**Since the founding of Devereux Schools in 1912, a philosophy of Special Education and therapeutic techniques of rehabilitation and care have been developed for the sole reason of insuring to the greatest extent the future happiness and well-being of the retarded or emotionally disturbed child. Many of these techniques originated at Devereux Schools. This series is launched during the Schools' 50th Anniversary in order that you may better understand the methods and goals of Devereux Schools.*

The Devereux Foundation

A nonprofit organization . . . Devon, Penna. • Santa Barbara, Calif. • Victoria, Texas
Helena T. Devereux, *Founder and Consultant* • Edward L. French, Ph.D., *President and Director*

Half a Century of Service to Children

PSYCHOLOGICAL TEST SPECIALISTS

Exclusive publishers and distributors of selected psychological tests

NEW!! NEW!! NEW!! NEW!!

Quick Test (QT)

A very quick and satisfactorily reliable screening test of intellectual ability (6 to 10 minutes total time for all three forms). Words are defined non-verbally by choice among drawings. Suitable for any person (no matter how physically disabled) who can signal a yes-no response in any way. Covers intelligence levels from two years to superior adult. Excellent for building rapport, and for quick testing before and after any procedure.
$8.00 for set of materials

Memory-For-Designs Test (MFD)

Simple drawing test of perceptual-motor coordination, serving as extremely sensitive detector of brain injury of many types.
$8.50 for set of materials

Group Personality Projective Test (GPPT)

Multiple-choice questions about ambiguous stick-figures make possible rapid and effective measurement of significant, factor-analytically isolated aspects of personality.
$13.50 for set of materials

Id-Ego-Superego Test (IES)

Subtests give independent estimates of personality components from a modified psychoanalytic point of view.
$28.50 for set of materials

Kahn Test of Symbol Arrangement (KTSA)

Uses symbolic plastic objects which are arranged and sorted; validly discriminates among normals, [illegible] brain-injured, and psychotics, and indicates accompanying personality dynamics.
$25.00 for set of materials

Full-Range Picture Vocabulary Test (FRPV)

Brief test of intelligence in two forms using the same 16 plates. Words are defined by choosing pictures. Valuable for intelligence screening, and with the physically handicapped.
$7.50 for set of materials

Geist Picture Interest Inventory (GPII)

Provides for more valid expression of interests by choices among alternative pictured activities. [illegible] qualitative and quantitative interpretations are possible.
$6.00 for set of materials

Proverbs Test (PT)

Group and individual forms determine similarity of thought processes to those of various normal and clinical groups. Highly sensitive indicator of psychotic processes.
$4.50 for kit

Vocational Apperception Test: Advanced Form (VAT:ADV)

Separate series of plates showing male and female college-related occupational activities, accompanied by explanation of use and interpretation as a projective test.
$7.25 for set of materials

Famous Sayings (FS)

Agreement with proverbs, aphorisms, and folk sayings is analyzed to determine personality structure. Particularly useful in personnel selection.
$4.50 for kit

It Scale For Children (ITSC)

An experimental picture scale for measuring sex role preference in children.
$15.00 for set of materials

Kahn Intelligence Tests: Experimental Form (KIT:EXP)

Relatively culture-free intelligence tests requiring no reading or writing. Supplementary scales for assessment of intelligence of blind, ability in concept formation, recall, and motor coordination. One scale can be administered by sign language.
$26.00 for set of materials

PSYCHOLOGICAL TEST SPECIALISTS

Box 1441 MISSOULA, MONTANA